

Free trade everywhere in chains

The periodic trade war of words between the United States and Japan is about to be resumed; but the United States should be worrying about Europe.

THE United States, more than ever alarmed that its industrial prowess is declining (the Office of Technology Assessment was saying as much only last week), seems bent on choosing the wrong way to put things right. So much can be inferred from the way in which the US administration is loudly advertising its impending negotiations with Japan on their mutual trading relationship. The objective seems to be to promise everybody, not just Japan, that there will be tough talking in the next four months, at the end of which the US government has to decide whether Japan is to be singled out for retaliatory trade sanctions under the terms of the 1988 Trade Act. (The European Community and Brazil were put in the same doghouse by the administration's finding, last year, that they were also guilty of unfair trade practices.) As a preliminary, Mr Toshiki Kaifu, the prime minister of Japan, appears to have been invited to last weekend's meeting with President George Bush, at Palm Springs in California, in terms so pressing that he may have judged it imprudent to decline.

Irony

There is a delicate irony in the timing of these events, which coincide with Japan's first announcement in six years that its balance of trade with the outside world has been negative for a single month. What seems to be happening is that there has been a genuine liberalization of the Japanese import trade (even though Liberal Democratic Party politics, reinforced by last month's re-election of the ruling party, mean that rice farmers will continue to be subsidized) and that Japanese consumers have prudently discovered an interest in many cheaper products made by people other than themselves. But there are wider reasons why the US administration, to its credit as convinced as ever that free trade is in everybody's interests, should seek to blunt the sharp edges of the trade policy it has inherited.

First, the trading relationship with Japan is not the only one that should matter to the United States. Indeed, the United States should now be worrying more about its relationship with the European Communities, now less than three years away from becoming the single market they have pretended to be for the past 30 years. Tough talking is also in fashion there. The long-standing complaint about Europe's protection of its farmers at a cost of \$30,000 million a year would be more persuasive if US

farmers were not supported (by a different mechanism) on the same scale. But by this means, the United States is asking no less than that the constitution of the European Communities should be changed. There is a case for doing that, but it will not be conceded in time to meet the timetable set by the Congress. In the past few days, the United States has also formally complained that the European telecommunications market is not as open as it should be to imports of capital equipment from elsewhere, but has not acknowledged the degree to which the European communications industry has lately been transformed, both by denationalization and technical advance. More guile would win more concessions, and might even nudge Europe away from the protectionism that threatens it.

Forum

Second, the United States seems to have forgotten that there is a multilateral forum in which these and other issues can be negotiated. For the past three years, the members of the General Agreement on Tariffs and Trade (GATT) have been haggling over the revision of this treaty, seeking such things as an extension of non-discriminatory rules that now apply from manufactured goods to services — banking and insurance, for example. The trouble with GATT's rules is that they are too numerical. Countries are allowed to levy tariffs on imported goods, for example, but are required not to have different tariffs for different suppliers unless one can be shown to be exporting at unrealistically low prices. (There are also agreements between the member states on the rates at which tariffs should apply.) But there is no reason why the same framework should not be used to negotiate the removal of the discontents that plague the United States. It would make for better tempers all round, and would probably be more efficient as well, if the United States were to look to GATT for a resolution of the non-tariff quarrels it has recently been picking with its trading partners. Even if that meant giving up the 'Buy America' provisions of the Walsh-Healey Act, the outcome could be worthwhile.

Misconception

The underlying misconception seems to be the conviction, especially in the US Congress, that industrial success somehow equates with the successful sale to others of a



range of specified technically advanced products. Telecommunications equipment and Earth satellites are high on the US list at present, but they will become increasingly symbolic of last decade's technology. So, eventually, will be even supercomputers, one of the bones of US contention with Japan, in respect of which the Japanese government seems on the point of increasing the funds available to Japanese universities so that they will no longer have to depend on hefty discounts from Fujitsu for the purchase of these machines, but may buy full-price Crays instead.

Most probably, the administration is aware of all this, but feels trapped by the pressure from the Congress. It is even possible that the tough talking at the outset of the negotiations is designed to relieve the pressure from the Congress, thus allowing the administration to reach a reasonable accommodation with Japan.

But that is a risky strategy. It would be safer if the government tried to persuade its disparate constituents of the more substantial truth that the virtues of free trade are so great that they do not require reciprocity. If somebody else can sell manufactured goods or services more cheaply than they can be provided domestically, they constitute a bargain that should be jumped at. These days, there are all too many other things for idle hands to do. □

Unacceptable risk

The reason why radiation may cause leukaemia in offspring should be pinned down as a matter of urgency.

THE managers of the British reprocessing plant for nuclear fuel at Sellafield have had a predictably torrid time in the past two weeks. The conclusion by Professor Martin Gardner and his colleagues that some cases of leukaemia in people under 25 may have been caused by their fathers' exposure to radiation at work (see *Nature* 343, 679; 1990) could only have made more sharply poignant the decisions that those who work with radiation have always to make: will the small but non-zero extra risk be identifiable above the noise of mortality from natural causes among those exposed and of unspecific genetic defect among later generations? The trouble, with the now-plausible hypothesis that radiation exposure may occasionally be responsible for leukaemia among immediate offspring, is that the noise has been reduced virtually to zero. For at least some time, each male radiation worker who fathers a leukaemic child will jump to only one conclusion. Will nuclear plants now be able to recruit workers, and how?

The important letter from Dr J. H. Dunster on page 98 bears on that practical question as well as on the broader questions that have now been raised. Dunster, who retired only a year ago as director of the British National Radiological Protection Board, and who is now chairman of one of the sub-commissions of the International Commission on Radiological Protection (ICRP), presents only back-of-the-envelope calculations. In the circumstances,

nothing more elaborate is possible or justifiable. Gardner's study, after all, makes it quite plain that the association noted at Sellafield does not prove that radiation exposure causes leukaemia. And if there is such a link, there is nothing as yet to show whether what matters is a person's exposure in the few months before conception of a child or the accumulated dose, from natural and artificial radiation, since birth.

With all these reservations, Dunster's arithmetic shows that, if short-term dose is what matters, there is a one in 400 chance that a child fathered by a man exposed to 20 mSv of radiation will contract leukaemia. Rightly, he says the risk is "small, but not negligible". If, on the other hand, lifetime dose is what matters, the risk that a child will be leukaemic will be roughly the same if its father has been exposed to 200 mSv. Moreover, with the uncertain numbers arising from the Gardner study, if lifetime dose determines whether a leukaemic child is fathered, people's exposure to natural radiation may be a sufficient explanation of the occurrence of all leukaemia in the young.

The most urgent need is to discover whether the Gardner effect (if not just chance) is caused by the short-term radiation dose or by that accumulated over a lifetime. Luckily, there is much that research can do. Already there is some evidence of a linkage in the general population between paternal age and the occurrence of leukaemia among children, which implies an influence for the accumulated dose of natural radiation. If this linkage could be made quantitative by more extensive investigations, it would be possible to estimate the influence of accumulated radiation dose to male germ cells. That would at least allow the managers of nuclear plants and their employees to know what might be gained by opting out of radiation exposure temporarily.

The wider issues are those of the general management of radiation exposure. Natural radiation provides everybody with an annual dose of about 1 mSv and the natural occurrence of leukaemia in the young (in Britain) is about 1 in 2,750. If there is a causal relationship between radiation exposure and the occurrence of leukaemia, and if the numbers are those suggested by Dunster's provisional calculation, it is difficult to believe that the managers of nuclear plants will in future be able to regulate their affairs by yardsticks even as carefully devised as the latest proposals of ICRP, now in circulation in draft form.

Instead, it will be necessary to arrange that every employment interview is followed immediately by a counselling session, that workers at nuclear plants are not merely informed but also consulted about their radiation records and that the general exposure to radiation is further reduced. It is not that radiation risks have been grossly underestimated in the past, but that the Gardner study points to a weakness in the definition of what risks are acceptable. The principle has been that the consequences of exposure should be so unlikely that they are lost in the noise. If radiation causes leukaemia in offspring, even with a likelihood as small as one in 400, the consequences will no longer seem inconspicuous to those concerned. □

Pollution as Czech public enemy number one

- Dangerous levels of sulphur dioxide
- Reduced average life expectancy

Prague

CAUGHT between ecological and economic crises, the new government of Czechoslovakia is working feverishly to find a plan to preserve what is left of the ravaged environment without ruining the economy.

The coal-mining regions of northern Bohemia have been devastated by the effects of strip-mining and the burning of high-sulphur coal. During temperature inversions, which normally last two to four weeks each winter, the sulphur dioxide level can reach 2,000 μg per mm^3 of air, more than 20 times levels considered harmful to health.

Until recently, most such data were kept a secret by the Communist government, but "you did not need measurements to know something was wrong", says Vaclav Slavik. "You couldn't see more than 2 metres" during an inversion, he says, because of the acrid yellow fog.

Life expectancy in northern Bohemia is 1.5 years lower than the Czechoslovak average and the chances of dying of cardiovascular disease or cancer are significantly higher, says Jan Kasal of the Regional Health Station in Usti nad Labem.

The forests across Czechoslovakia are also endangered by the pollution. Svatoimir Mlcoch of the Czech Environment Ministry reports that 40–60 per cent of the forests in Bohemia and Moravia are dead or severely damaged by pollution.

The situation is no better in Prague. Czechoslovak television reported in mid-February that the level of nitrogen oxides in central Prague, which supposedly never exceeded a limit set by the old government at 100 μg per mm^3 had reached 4,100 μg per mm^3 .

"We have all the political support we need" to enact anti-pollution measures says Bedrich Moldan, the harried new minister for the Czech lands, who is responsible for the environment in about two-thirds of Czechoslovakia. A poll released last week on Czechoslovak television showed that 83 per cent of respondents considered the environment the "first priority" for the new government, followed by health care. (An improvement in the variety of goods available in stores came tenth.) Another recent poll showed that 43 per cent of Czechoslovak citizens would accept a cut in their standard of living if the money were invested in improving the long-term welfare of the nation, its economy as well as its environment.

But despite the high level of public support, Moldan, an analytical chemist specializing in acid rain who succeeded in publishing a few data about the environment even under the old regime, says that a system to deal with pollution is still lacking. In the short term, he says, it will be very hard to reorganize the Czechoslovak economy along sound environmental principles.

Greater freedom may even make control of pollution more difficult. It will be impractical, he says, to grant autonomy to factories and at the same time expect them to install filters or water treatment plants to reduce pollution, because they may just decide to invest their money in something else.

Czechoslovakia is enjoying an exceptionally mild winter this year with fewer smog alerts than usual. But Moldan says he is afraid of strikes and demonstrations in northwest Bohemia if the pollution returns to its normal level. "The people are willing to make sacrifices", he says, "but they don't want to breathe polluted air any more." Pollution problems may even encourage harmful political battles.

The Ecoforum group, a part of the Civic Forum that organized the peaceful 1989 revolution, has sent a letter to Moldan and other leaders blaming the disastrous economic conditions in northwest

Bohemia on the need to "secure the supply of goods and energy to other people in Czechoslovakia". Moldan says it is politically very dangerous to blame the problems on other citizens instead of on the previous government.

The first priority, Moldan says, is a comprehensive energy policy that includes restructuring the economy around light industry in place of steel and other heavy industries.

Czechoslovakia, like other Eastern European countries, is caught between two unattractive energy alternatives — burning coal with a high sulphur content or building more nuclear power plants.

Even coal miners would favour nuclear power (despite the potential loss of their jobs), says Moldan, and the people who live near the two existing nuclear plants and the two under construction favour other energy sources. Moldan advocates completing the plants now under construction but phasing out nuclear energy in 20 years.

Moldan cannot even shut down the worst polluters in the country until there are alternatives. The factory at Ostrava which recycles used oil is a case in point. The factory is "old and dirty", he says, but if it is shut, "people will just dump the oil on the ground".

One obvious strategy — getting help from outside — will not be easy to apply after President Vaclav Havel declaration that Czechoslovakia wants to accumulate no new foreign debt. But Moldan will visit West Germany later this month to see about accepting help to reduce the sulphur content in the smoke from coal-fired power plants.

Steven Dickman

SPACE TRANSPORT

Firing pennies to heaven

Washington

FACED with the soaring cost of satellites, a US congressional agency is proposing that 'lightsats', 'fatsats' and legions of tiny disposable rockets with ice as fuel could provide imaginative ways to save money.

A new report from the Office of Technology Assessment (OTA) says that although launching a pound of payload into low-Earth orbit currently costs about \$3,000, the payloads themselves cost between \$200,000 and \$800,000 per pound. Using cheaper rockets — even an order of magnitude less expensive than current designs — would cut total cost by only a few per cent, OTA says.

A more effective way of making savings is to allow satellites to be larger, heavier and less technologically refined ('fatsats'), or simpler and lighter ('lightsats'). OTA's most startling suggestion is that simple tube-like rockets filled with ice or plastic could carry 20 kg 'microsatellites' into

space. Power would come from a ground-based laser that would be directed at the inert fuel and vaporize it. The launch cost would be about \$120 per pound. At up to 60,000 launches a year, such a systems could put twice as much tonnage — in 20 kg lumps — in space as current US programmes.

But the plan would require 25 MW lasers, about five times as powerful as the best that have been so far demonstrated. Developing such a laser could cost \$525 million, says OTA, quoting figures from the Pentagon's Strategic Defense Initiative.

And still unanswered is how useful 60,000 small packages in space might be. Fuel for interplanetary missions is one possibility, but even that proposal has its sceptics. "Think of all the astronauts you'd need to decant it", says Federation of American Scientists space analyst John Pike. G. Christopher Anderson

New computer link for Japan

Mishima

THE DNA Data Bank of Japan (DDBJ) at the National Institute of Genetics in Mishima will this month open a high-capacity computer link that will allow exchange of much larger volumes of data with DNA databanks in Europe and the United States. But Japan's contribution to international efforts to handle the growing flood of DNA data remains pitifully small.

At a time when GenBank in the United States and the EMBL Data Library each employ teams of tens of researchers to run their databanks, Japan still struggles along with only two full-time researchers and a couple of part-timers at DDBJ. Not surprisingly, DDBJ at present processes only about 3 per cent of all the DNA sequences published worldwide; the remaining 97 per cent are dealt with at EMBL and GenBank.

DDJB constantly exchanges data with EMBL and GenBank and the volume of traffic has grown to over 1 million bits per day. This volume is expected to grow at least five times when the databanks move to a new relational database system in the near future, according to Sanzo Miyazawa of DDBJ. Hence the need for the high-capacity computer link.

The new computer line, which will be connected on 28 March and can carry 64 kilobits per second, is between DDBJ and the University of Tokyo.

Last year, the university opened Japan's first high-capacity computer link with the outside world, the Todai International Science Network, with a leased line to the University of Hawaii (see *Nature* 340, 670; 1989). DDBJ will gain access to scientific networks in the United States and Europe and thence to the GenBank and EMBL databases through Tokyo University. The new link will also allow DDJB to form a distributed database with DNA databanks at several locations in Japan.

The three databanks have agreed that from now on DDBJ will process all DNA sequences produced in Japan. This will increase DDBJ's data share to about 10 per cent. But the bank is in desperate need of more manpower to handle the increased workload. The rigid employment policies of the Ministry of Education, Science and Culture (MESC), to which DDBJ is affiliated, make it very difficult for the bank to employ new staff. DDBJ has won permission to create a new position for an assistant professor but more temporary staff are also needed for annotating DNA sequences as they are input.

The new leased computer line will also be a considerable drain on DDBJ's very limited budget. Despite privatization of the telecommunications industry several years ago, Japan's domestic telecommuni-

cation charges remain very high compared with those of other developed nations. The charges for leasing the line to Tokyo University will be nearly ¥5 million (\$33,000) a year, or 50 per cent of the bank's annual grant. But, Miyazawa expects to get a new grant from MESC to cover these extra costs.

The lack of IBM-compatible computers in Japan also adds to DDBJ's headaches. Japan is the only developed nation where IBM-compatible personal computers are not the standard; rather, the market is dominated by NEC computers. GenBank recently developed the new software 'authorin' which allows scientists to input DNA sequences directly to databanks, thereby easing the workload for the databanks. But the software is of virtually no use in Japan until Miyazawa rewrites it into NEC-compatible form, something he hopes to do this summer. In any case, few Japanese biologists are computer-literate and it is likely to be a long time before large amounts of data are submitted in this way.

Japan also lags behind GenBank and the EMBL data library in the use of CD-

ROM to distribute DNA data to scientists. EMBL began using CD-ROM last September and GenBank began this month. But DDBJ continues to use magnetic tapes, which are cumbersome and hold fewer data, because there is little demand for such a service from Japanese scientists and the tapes are easier to copy, Miyazawa says.

Japanese researchers promoting the human genome project are considering Mishima as a possible site for a new bioinformation centre to handle data arising from the project. But although it is "quite possible" that the centre will be located at the National Institute of Genetics, it will probably be separate from DDBJ, according to Takeshi Seno of the institute.

Miyazawa says that DDJB simply does not have the capability to cope with the massive amount of data that is expected to arise from the human genome project. Some members of the institute fear that the centre will be a "service-orientated" facility providing information to researchers throughout Japan and their duties for the centre will cut into valuable research time. As it is, Miyazawa and his assistant have very little time for their own research because of commitments to DDBJ.

David Swinbanks

HEALTH RISKS

DoE to give up studies?

Washington

AN independent panel assembled to advise the Department of Energy (DoE) on its troubled epidemiology programme is expected to report this month that long-term studies on the effects of radiation and chemical hazards at DoE facilities cannot be credibly done by the department.

Kristine Gebbie, the panel's chair and Washington state's senior health official, says that the panel's draft report will recommend that DoE provide financial support for another agency such as the Department of Health and Human Services (HHS) to carry out long-term epidemiological studies.

"DoE's credibility has been attacked by so many people", Gebbie says. The agency "has to break down the overwhelming perception that it is a closed shop". Reforming the programme is one element of DoE director Admiral James Watkins' 'ten-point plan' to rebuild the agency. DoE has been accused of distorting and suppressing the results of epidemiological studies and of putting pressure on researchers who study the health of its workers.

At an advisory panel hearing in Columbia, South Carolina, last month, Gregg Wilkenson, a Los Alamos National Laboratory researcher, testified that DoE officials had put pressure on him to rewrite a report showing that workers at

DoE's Rocky Flats plutonium processing plant near Denver, Colorado, had higher than average cancer rates. Wilkenson said that in 1986 and 1987, while he was the laboratory's epidemiology group leader, he was asked to make the report appear more favourable to DoE.

Watkins has called for an investigation of Wilkenson's charges. "I am disturbed by even the allegation of DoE interference. I feel strongly that the independence of scientists to do their research and publish without departmental interference is critical to the credibility of their research and findings."

But Los Alamos spokesman John Webster says the laboratory has investigated the allegations and found no evidence of wrongdoing. "There was no attempt to berate Dr. Wilkenson", Webster says. "The study was published as written." Congressional critics are unlikely to be satisfied with the changes DoE is planning for its epidemiology programme. Senator Tim Wirth (Democrat, Colorado) plans to continue a legislative effort to move the entire effort — management and all — out of the energy agency and into HHS. His aides say that he is concerned that if the heart of the programme is allowed to stay in DoE, it may revert to secrecy and censorship once Watkins has left the agency.

G. Christopher Anderson

SPACE STATION

Redesign plans assailed

Washington

THE space station last week came under new attack from lawmakers who fear that it is becoming more bus station than scientific laboratory.

At a hearing to examine the 1991 budget for the National Aeronautics and Space Administration (NASA), legislators warned NASA that changing the project's design to accommodate the president's proposal to send humans to the Moon and Mars risked breaking research commitments and alienating international partners.

"We were sold the space station partly on its ability to do microgravity research", Representative Bill Green (Republican, New York) told the President's science adviser, D. Allan Bromley. If the space station is to be used as a "transportation node" for interplanetary travel or orbital missions, Green pointed out, the jostling of men and machines on board could make low-gravity science difficult.

"What we have now is a space station that we are struggling to keep as a scientific laboratory", added Representative Bob Traxler (Democrat, Michigan), chairman of the appropriations subcommittee that funds NASA. "Our international partners' first area of concern is microgravity research. That should send us a signal." Some current NASA scenarios would allow for as little as two years of high-quality microgravity time before the station is adapted to serve as a construction and servicing hub for other proposed missions. Ian Pryke, head of the Washington office of the European Space Agency (ESA), confirms that basic science is indeed ESA's top priority, but that it is not inconceivable that the agency

WHITEHEAD INSTITUTE

Fink appointed director

Boston

THE Whitehead Institute for Biomedical Research in Cambridge, Massachusetts, has named Gerald R. Fink as its new director. Fink, who will succeed David Baltimore in the post on 1 July, is currently American Cancer Society professor of genetics at the Whitehead Institute and in the department of biology at Massachusetts Institute of Technology. One of the original members of the Whitehead Institute, Fink pioneered the use of yeast as a model for genetic research, and was the first to develop a method to introduce foreign genes into yeast. An outspoken proponent of increased support for research in plant biology, Fink is a member of the National Academy of Sciences and the American Academy of Arts and Sciences as well as a former president of the Genetics Society of America.

Seth Shulman

could want a part in a Moon/Mars project as well.

"We entered into a partnership to build an orbiting laboratory", Pryke says. But he says that ESA, which along with Canada and Japan represents most of the international component of the project, is waiting to see how plans for a Moon/Mars project develop before it decides whether to ask NASA to guarantee adequate resources for science.

One possible solution is to relegate microgravity science to smaller modules that orbit near the station but are not physically connected to it. The ESA Man-Tended Free Flyer is one such module and is already scheduled for launch in the late 1990s. NASA spokesman Mark Hess says the space agency is considering further separation of construction and science on the space station, a redesign that could change the project's concept from that of a single structure to an orbiting swarm of smaller modules.

MARS INITIATIVE

NASA's plans too timid

Washington

THE US National Research Council (NRC), in a report published last week, gives a general endorsement to plans put forward by the National Aeronautics and Space Administration (NASA) for accomplishing the voyage to Mars, but also adds its voice, if rather discreetly, to the chorus of critics who believe that NASA is being too timid in its thinking. The NRC suggests that establishing a space station in orbit around Mars would be an easier and safer first goal than going directly to the surface, and argues that too little attention is being paid to human issues, particularly the long-term effects of weightlessness and the possibility of reducing these by using artificial gravity.

The NRC report, produced at the request of Vice-President Dan Quayle in his capacity as chairman of the National Space Council, is a review of NASA's '90-Day Study', released last November, in which five strategies for establishing a human presence on Mars were sketched out (see *Nature* 342, 607; 1989). In all of those plans, after a 'practice' base has been set up on the Moon, a direct journey is undertaken from Earth orbit to the martian surface. The NRC report, however, points out that the descent to the martian surface is likely to be the hardest part of the trip, and recommends that a station in Mars orbit would provide a safe staging post from which the final assault could be mounted. The report chides NASA for placing too much reliance on the feasibility of 'aerobraking', by which

But even traditional congressional supporters of the project are growing wary of NASA's willingness to change its basic configuration seemingly at will. "It troubles me that [these issues] are being



looked at six years after [the space station] was approved. We're building it, then looking for a purpose for it", Green said last week. Traxler warned Bromley that the eleventh-hour redesigns are "destroying [the administration's] credibility and harming science".

G. Christopher Anderson

spacecraft would descend to Mars beneath vast parachutes.

But the NRC review committee gives little encouragement to the idea, promoted by physicists Edward Teller and Lowell Wood, that a cheap, quick trip to Mars could be made by 'space inflatables' (see *Nature* 343, 301; 1989). The report observes laconically that the difficulties of turning the idea into practical technology are "underestimated".

David Lindley

■ Under pressure from Congress, NASA last week finally released the cost estimates to go with its four-month old '90-Day Study'. NASA expects that the trip to Mars will cost \$541,000 million (in 1991 dollars) over 30 years, which is \$241,000 million more than administration officials gave late last year.

To establish a permanently manned base on the Moon by 2001, \$100,000 million will be needed, and maintaining it until 2025 could cost \$208,000 million. Visiting Mars twice, culminating in a 600-day stay in 2018, would take another \$233,000 million. If the lunar base is not permanently manned, and if only three short visits to Mars are made, the cost falls to \$471 million.

White House science adviser D. Allan Bromley quickly warned Congress to take the figures with a grain of salt. The '90-Day Study' from which the cost estimates came was "a very crude" process, Bromley told legislators on the House appropriations committee. "I wouldn't put any credibility in the numbers that came out of that", he said.

G. Christopher Anderson

Blind eye for financial perils

London

SIR Peter Swinnerton-Dyer, chief executive of the UK Universities Funding Council (UFC), last week surprised a House of Commons committee when he admitted that he had no means of knowing whether a financial crisis similar to that which almost forced the closure of University College, Cardiff in 1987 was building up at another British university.

Giving evidence to the Public Accounts Committee, Swinnerton-Dyer said the UFC has to rely on the accuracy of financial forecasts provided by individual universities. The UFC "is not staffed or funded", he said, to check these figures independently. Accurate financial forecasting by universities beyond 1991 had also become more difficult since the UFC introduced its competitive bidding system for undergraduate places, Swinnerton-Dyer said.

JAPANESE SCIENCE

Polite bows to the West

Tokyo

JAPANESE scientists have a low opinion of the quality of Japanese research institutes and universities according to a survey released by the Science and Technology Agency (STA) last week.

Only a handful of the scientists surveyed were prepared to say that Japan has any centres of academic excellence; instead, nearly all chose institutes in the United States and Europe when asked to name such centres.

The respondents were 385 members of the Science Council of Japan, an elected body of distinguished academics. They were asked to name up to three centres of excellence overseas and one or more in Japan if they thought any existed. Only 11 scientists, less than three per cent, named any Japanese institutes at all.

The National Institutes of Health in the United States received the most votes (38), followed by Massachusetts Institute of Technology (29) and Max-Planck-Gesellschaft (28) in West Germany.

The National Laboratory of High Energy Physics in Tsukuba was one of the few Japanese institutes to win some votes (9).

But are these scientists doing their own country justice or are they just being modest? Japan's Institute of Space and Astronautical Science (ISAS), which last month launched a spacecraft to the Moon, is widely recognized outside Japan as a world leader in X-ray astronomy and other areas of astrophysics.

It is also one of the most international institutes in Japan with a large complement of foreign scientists. But this institute received only two votes. David Swinbanks

Universities will not know until next year how many students they will be teaching in each subject from 1991-92.

UFC's written evidence to the committee shows that 44 out of 74 universities (including colleges within the University of London) predict a deficit by the end of the present academic year.

Swinnerton-Dyer identified "a lack of realism" amongst the management of some universities as a contributing factor, but refused to comment in a public meeting on the financial position of particular institutions. Doing so might alarm universities' creditors, he argued.

Committee members were critical of UFC's negotiations with the University of London, whose finances are causing the UFC "serious anxiety". Swinnerton-Dyer said he had ensured that London's funds were properly managed so far as the autonomy of London's colleges allowed. Terry Davis, a Labour Member of Parliament, accused Swinnerton-Dyer of "sheltering behind academic freedom".

The committee was concerned that in reducing deficits, institutions might have to make such large cut-backs that they no longer merited the title of 'university'. Swinnerton-Dyer said that this had not yet been the case, but "enough cuts would make this problematical".

The Committee of Vice-Chancellors and Principals (CVCP) is also worried about the maintenance of standards in the face of accumulating financial deficits. In a statement delivered to the Public Accounts Committee to coincide with its inquiry, the CVCP warns that "unless more money is provided . . . the present and planned expansion of student numbers in universities will mean a decline in the quality of education provided".

Swinnerton-Dyer also revealed that the UFC may soon face court action from the University of Edinburgh. The UFC is withholding payment of about £200,000 to cover pension payments to Edinburgh academics who took early retirement under the old University Grants Committee's (UGC) restructuring programmes, on the grounds that these staff have been re-engaged on a part-time basis. But Sir David Smith, principal of Edinburgh, says that the UFC has changed the rules. In any case, Edinburgh argues, the staff concerned are not being paid from the UFC's block grant.

Swinnerton-Dyer said that the UGC never intended that staff could be re-engaged in this way, except to fulfil existing course requirements, but conceded that UGC circular letters explaining the early retirement programmes may have been ambiguous.

Peter Aldhous

Gamma-ray detector to be built

London

THE British Science and Engineering Research Council (SERC) and the French Centre National de la Recherche Scientifique (CNRS) last week signed a £5 million agreement to build EUROGAM, the world's most sensitive gamma-ray detector array.

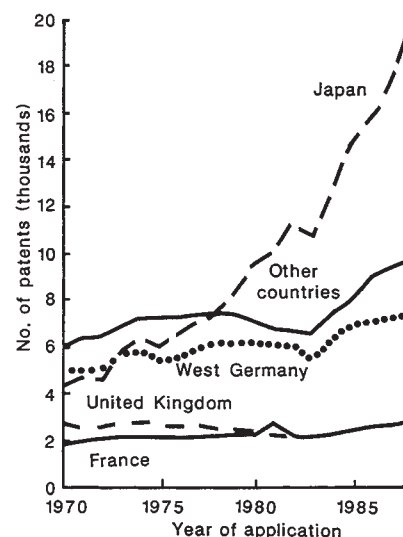
The equipment will investigate the structure of atomic nuclei, produced by ion collisions. EUROGAM will be built at SERC's Daresbury laboratory, in Cheshire, coming on line in late 1991. A year later, the entire array will move to the CNRS accelerator in Strasbourg.

Professor E. W. J. (Bill) Mitchell, SERC chairman, describes the sharing of equipment between sites in different European countries as "an interesting and new way of working". Although the arrangement is chiefly a diplomatic measure, necessary to attract joint support, there may be some scientific benefits: Daresbury and Strasbourg have complementary instrumentation. Researchers at Daresbury initially sought full funding from SERC, but were asked to find a collaborator. Peter Aldhous

US patent awards

Washington

NEARLY half of all US patents are now awarded to foreigners, according to figures in the new issue of the US National Science Board's biennial report, *Science and Engineering Indicators—1989*. That percentage fits quite well with figures showing that the United States does almost half of the Western world's total



research and development. Japan's burgeoning patent inventory also fits with its rapidly increasing support for research and development. Between 1974 and 1984, the number of patents Japan gained in the United States increased around fourfold; its expenditure on research and development increased threefold. The UK share of US patents remained almost unchanged over the same period, as did its research and development expenditure in constant dollars.

Alun Anderson

AIDS MEETING

WHO to aid Eastern Europe

Munich

FEARS that a major epidemic of AIDS could occur in Eastern Europe has stimulated the World Health Organization (WHO) to join with national governments in a new preventive programme. The US government is making a grant of \$1.5 million but more will be needed.

Although there are very few AIDS cases in Eastern European countries — a total of just 306 was reported as at 31 December 1989 in seven Eastern European countries plus the Soviet Union — WHO epidemiologist David Heymann said that "the potential is probably there" for a rapid spread of the disease.

In February, 550 children in Romania were found to be infected with HIV (human immunodeficiency virus), the virus that causes AIDS (see *Nature* 343, 579; 1990). Many will die before they reach the age of three. They are thought to have been infected through the use of unsterile needles used during 'micro-transfusions' of blood and blood products.

At a meeting in Copenhagen last week, Eastern European countries for the first time reported the number of people infected with HIV (see chart). Bulgaria reported 154 known HIV-positive people, the Soviet Union 899 and Yugoslavia 1,713. These countries estimate that up to three or four times as many people are really infected. The largest estimates are from countries where drug abuse is known to be a problem.

CONSERVATION

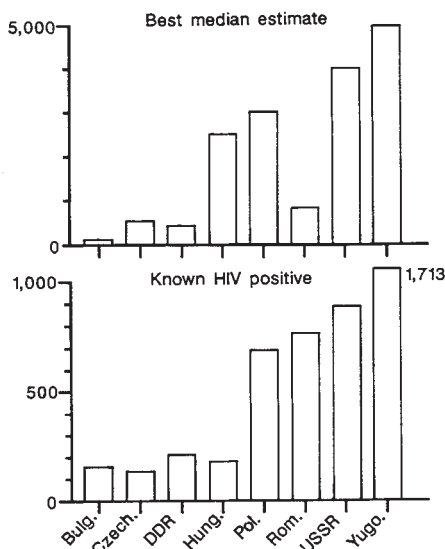
British agencies need strong scientific base

London

THE House of Lords Science and Technology committee has urged the British government to spend more to ensure that the separate English, Welsh and Scottish conservation agencies that will replace the Nature Conservancy Council (NCC) each have a strong scientific base. But the three agencies should not aim to be scientifically self-sufficient, the committee's report says. The joint committee that is to oversee the agencies should ensure cooperation and prevent duplication.

The report does not comment on the merits of the widely criticized decision to devolve the NCC (see *Nature* 343, 4 & 15; 4 January 1990), but says the government's Environment Protection Bill, which proposed the change, "shows signs of hasty drafting". In particular, the committee wants the bill amended so that the joint committee can make legally binding decisions in the event of disputes between agencies.

Peter Aldhous



AIDS in Eastern Europe. Data courtesy WHO.

At the invitation of the governments, WHO will help set up AIDS education, testing and monitoring programmes in all Eastern European countries except Albania, which claims that it has no cases of AIDS. WHO will help each country to formulate an 'action plan' against AIDS after assessing the spread of the disease and the steps already being taken to prevent transmission of the virus.

WHO activities will include an Eastern European workshop on education and public information in Dresden in May and an exhibition on AIDS and youth in Warsaw in June. All countries reported that they were screening "100 per cent" of blood and blood products used in transfusions. Heymann said that some may be using home manufactured test kits. He stressed that the kits must be tested to be sure that they meet international standards. Tests that give false negative results could be very harmful, he said.

After the evaluations are complete, the next step will be to begin testing and reporting of AIDS cases. WHO will help monitor the spread of AIDS by testing blood samples taken from unnamed individuals for other medical reasons. WHO will convene a meeting of Western governments and private donors in August to discuss concrete help for individual Eastern European countries.

WHO officials welcome the change in attitude seen in Eastern European countries which have previously denied that there was an AIDS problem. Heymann said in the past six months, mandatory testing of risk groups had been abandoned in favour of widespread voluntary testing and counselling. WHO this week sent an epidemiologist to Romania to help the government trace the origins of the AIDS outbreak there.

Steven Dickman

AIDS DRUG TESTS

Placebo or no placebo?

London

WORLDWIDE concern that experimental drugs for the treatment of AIDS may be withheld from patients while still under test has led British and French researchers to take a novel approach to the design of clinical trials for dideoxyinosine (ddI). The drug may help to prolong the lives of people with AIDS.

Tony Pinching, from St Mary's Hospital, London, who will run the British trial, says that unlike US pilot studies on ddI, his trial will allow patients to choose whether or not they join an experimental group which includes a placebo treatment. One group (option A), will be a standard randomized trial, with patients allocated across a placebo group, and low- and high-dose ddI treatments. Option B will exclude the placebo, so patients can be sure that they are receiving ddI.

US ddI trials have demanded a strict scientific protocol, including placebos. But pressure from AIDS activists led to widespread 'compassionate release' of the drug by its manufacturers, Bristol-Myers Squibb, under the Food and Drug Administration's 'treatment IND' (investigational new drug) measure (see *Nature* 340, 331; 1989). Pinching estimates that about ten times more people have received ddI in this way than through the better-monitored experimental trials.

The Medical Research Council (MRC), which is funding the British trial, says that the design will ensure that "all eligible patients can be entered into a formal study". Although there were doubts about whether many AIDS patients would volunteer for the option that included a placebo, Pinching says that, after consulting researchers, "some patients are already booked into option A".

The only drug at present licensed in Britain for treatment of AIDS is Wellcome's zidovudine (AZT), but 30–40 per cent of patients have to come off AZT within a year of treatment because of its toxic side-effects (including nausea, anorexia and muscle wasting). The new trial will look at the effects of ddI in several hundred of these AZT-intolerant patients.

Like AZT, ddI inhibits the enzyme reverse transcriptase, to slow replication of the human immunodeficiency virus (HIV). US trials have shown that ddI decreases the amount of HIV antigen circulating in the blood of AIDS patients, but there is no firm evidence yet of beneficial clinical effects comparable to those of AZT. The new trials will also look for toxic side-effects of ddI — some numbness and inflammation of the pancreas was reported in the US trials.

Peter Aldhous

New laws divide lawmakers

Washington

US CONGRESSIONAL moves to increase penalties for animal-rights break-ins have run into opposition from the Justice Department.

Several congressional bills now under consideration make crimes by animal activists against research laboratories a federal offence, with harsh penalties and probable investigation by the Federal Bureau of Investigation (FBI).

But the Justice Department opposes the bills, not because of ideological disagreement over whether activist crimes should be pursued with more fervour, but because the Republican party objects to the transfer of power from the states to the federal government (see *Nature* 343, 580; 15 February 1990).

The White House Office of Science and Technology Policy (OSTP) is, however, strongly in support of the bills. In the letter sent to the House Agriculture Committee, which held hearings on the issue last week, OSTP director D. Allan Bromley and associate director for life science James Wyngaarden (the former director of the National Institutes of Health) said that they object to the Justice Department's policy and believe that the administration should support the legislation.

"The animal rights movement is a national movement, with considerable circumstantial evidence of conspiratorial behavior", the letter states. The letter advocates the use of the FBI to investigate "this growing, well-organized, well-financed, unscrupulous movement". It also notes that People for the Ethical Treatment of Animals (PETA), a prominent animal-rights organization, often has videotapes and issues press releases shortly after break-ins attributed to the Animal Liberation Front (ALF), an underground activist group.

PETA co-founder Alex Pacheco says his organization abhors violence. At a recent press conference, Pacheco confirmed that PETA can legally use material collected during ALF raids, but "we don't know who [the ALF members] are and we don't want to know", he said. Until legislation is passed to allow outside inspection of animal research laboratories, PETA would continue to use the ALF-supplied information in the same way as a journalist would use leaked information, he said.

In the latest incident, PETA last week confirmed that it had received copies of documents stolen by the ALF last month from the offices of Adrian Morrison, a University of Pennsylvania sleep researcher. Morrison has been the chairman of the society for Neurosciences' Committee on Animal Research since 1987, and is a vocal defender of animal research.

The differences in opinion between OSTP and the Justice Department over the best legislation measures to deal with laboratory break-ins are compounded by divisions within Congress itself, where there are now nearly 80 current bills that would create legislation either in support of or in opposition to animal-rights issues, according to the National Alliance for Animal Legislation (NAAL).

Fifty-seven of the active bills support an increase in animal rights, welfare or protection and have the support of NAAL. The other 21 bills propose additional

BRAZIL

Ecologist appointed to environment post

São Paulo

IN a move designed to demonstrate to the world Brazil's concern for the environment, president-elect Fernando Collor de Mello has appointed the ecologist José Lutzenberger as head of a new National Secretariat for the Environment.

Lutzenberger is one of Brazil's best known and most vociferous environmentalists. He has been a critic of government subsidy schemes that have led to the destruction of rain forest by encouraging cattle ranching, and he opposes the extension of route BR-364 through Acre to Peru. When BR-364 was built into Rondonia, it provided access for huge numbers of set-

tlements who rapidly turned the state into the most deforested region of the Amazon. With extension of the road, the same fate is likely for Acre.

Lutzenberger is also a defendant of debt-for-nature swaps and has attacked President José Sarney's refusal to link negotiation of foreign debt repayments and environmental protection measures. But it is not known if the new president will accept Lutzenberger's views on this issue.

The agency that Lutzenberger will head is a new one and will be created in a planned reform of the administration. Collor will assume power next week.

G. Christopher Anderson

Ricardo Bonalume

NUCLEAR POWER

Embattled Seabrook wins licence at last

Boston

NEW Hampshire's Seabrook nuclear power plant, perhaps the most bitterly opposed reactor ever to be completed, received a full-power operating licence last week from the US Nuclear Regulatory Commission (NRC). The licence comes after more than a decade of opposition and delays during which time construction costs for the plant rose to \$6,450 million, more than ten times the original estimate.

Seabrook opponents, including leading legislators in the neighbouring state of Massachusetts, say they will appeal against the NRC decision in court. Calling the licensing decision "unconscionable", Massachusetts Governor Michael Dukakis says that his state will base part of its legal challenge upon the fact that the NRC changed its own rules on three separate occasions to allow the controversial plant to proceed.

Because of its chequered history, Massachusetts Attorney General James Shannon claims that Seabrook has the "most legally vulnerable license the NRC has ever issued". Nonetheless, both sides in the dispute acknowledge that there is

little precedent for court intervention at this stage of the licensing process. So most observers predict that the plant is likely to be operating at full power later this spring.

Utility representatives at the New England Power Pool, which manages the region's electric power grid, say Seabrook's additional 1,150 MW of electricity will help them to combat the power shortages that have forced frequent voltage reductions in the region over the past year. Most consumers, however, are wary of the increases in utility rates that are likely to be passed on to them once the plant is operating, increases that are estimated to add an additional \$24 annually to the bill of every electricity consumer in New England.

Utility and nuclear industry representatives rightly declare the licence a victory for the embattled Seabrook plant, but participants on both sides of the dispute agree that the construction and licensing process has been so gruelling and costly that it may discourage the nuclear industry for a long time to come.

Seth Shulman

CONFLICT OF INTEREST

Fierce debate at Harvard

Boston

At a faculty meeting said to have been the best-attended and most tumultuous in recent memory, some 300 members of the Harvard Medical School last week debated a proposal for new conflict-of-interest rules that would bar clinical researchers from financial ties with companies that could benefit from their research.

The proposed new rules were developed over the past year by a blue-ribbon committee appointed by Daniel C. Tosteson, dean of the medical school. The panel was set up shortly after a widely publicized case in which a former Harvard researcher was alleged to have failed to disclose properly a financial interest in a drug he was testing (see *Nature* 335, 754; 1988).

The proposed new rules would forbid any Harvard faculty member involved in human studies from acting as a paid consultant to any company with a financial stake in his or her research, and would forbid clinical researchers from owning stock in such companies. Non-clinical researchers would be required to disclose potential financial conflicts but would be allowed to retain financial equity in companies that sponsor their research — provided that the sponsored-research was funded through university channels.

Although the proposed rules remain several steps away from adoption, they have sparked vociferous debate among Harvard Medical School faculty, who form a disparate group of researchers working in a variety of capacities at the school and at more than a dozen affiliated hospitals.

Barbara J. McNeil, head of the school's department of health care policy and chair

of the committee that developed the rules, says that her panel stands firmly behind the proposal. "The committee believes quite strongly that we have produced a good document, and one that reflects a changing relationship between industry and university researchers", she says. The proposed rules address "increasing interest" on the part of the public "to know what we [researchers] are doing", McNeil says, while mandating only what she terms "parsimonious, relevant disclosure".

Other researchers however, claim that the guidelines "throw the baby out with the bathwater" and, if adopted, will make many researchers leave Harvard for more lenient institutions.

Opposition to the new rules is particularly strong and vocal at the Harvard-affiliated Massachusetts General Hospital.

Participants on both sides claim that despite an outcry from the biomedical community that forced the National Institutes of Health to back down temporarily from new stringent conflict-of-interest rules (see *Nature* 343, 104; 11 January 1990), a change at Harvard is likely before long. At the minimum, guidelines for disclosure of researchers' financial ties are likely to be strengthened.

How stringent the new rules will ultimately be, however, remains uncertain. As John T. Potts, chair of general medical services at Massachusetts General Hospital explains, "Some people are belligerently entrepreneurial, and that won't do. Others find dealing with industry dirty, a position that would be disastrous for science and for our society. The trick is to get it right in the middle."

Seth Shulman

BRAZILIAN SPACE PROGRAMME

Blast off for big ambitions

São Paulo

With the launching of a 4.5 metre-long sounding rocket, Brazil last month inaugurated its new space centre, situated near the small town of Alcântara, in the northeastern state of Maranhão. The launch was attended by outgoing President José Sarney. \$215 million is being spent on the new centre.

The sounding rocket was launched to test the base's tracking radars. According to the Air Force, everything went well with the Sonda-II rocket which went to a 110-km apogee before plunging into the Atlantic. The Sonda family of rockets provide the technological base for the development of the the Veículo Lancador de Satélites (VLS) — the locally manufactured satellite vehicle launcher.

The future launcher will be a four-stage

rocket, 18.8 metres high with a 50 ton mass at launch, and capable of placing a small satellite — 100–200 kg — in a circular orbit, from 250–1,000 km.

The launcher's first goal is to put into space the Brazilian Data Collecting Satellite (SCD-1), under construction at the civilian Institute for Space Research. But the date of the launch is by no means certain; it has already been postponed from 1989 to 1992, and further delays are likely in the next few years.

The programme has been handicapped by the unwillingness of the United States to transfer rocket technology on the grounds that the launcher programme, which is the responsibility of the Air Force, could be diverted to the production of a ballistic missile.

Ricardo Bonalume

ENERGY CLEAN-UP

Tiger Team lands at Livermore

San Francisco

As part of efforts by the US Department of Energy (DoE) to restore confidence in environmental protection at its production, research and testing facilities, members of the agency's renowned 'Tiger Team' descended on Lawrence Livermore National Laboratory last week to assess the status of more than 30 separate facilities — including Site 300, where conventional explosives used to trigger nuclear warheads are tested.

About 50 Tiger Team members will take part in the four-to-six-week investigation, reporting on compliance with a host of federal, state and local regulations. The sprawling laboratory complex has close to 500 separate facilities. But to make its task manageable, the Tiger Team is concentrating on sections designated as either high or moderate hazard, including areas housing plutonium, tritium, or heavy elements.

In announcing the review, Tiger Team leader Ed Cumesty said the action is part of Energy Secretary James Watkins' efforts "to restore credibility to the department".

In a separate move, laboratory officials announced that pending DoE approval they were dropping plans to build a giant incinerator to burn several thousand tons of radioactive and toxic waste produced annually at the site. The incinerator formed a key part of a proposed \$45 million waste treatment facility. Instead, the laboratory is to focus resources on other means of waste disposal and decontamination.

Local environmental groups praise both the incinerator decision and the Tiger Team review but have repeated their demand for an independent oversight body at the laboratory. Widespread ground-water contamination has been discovered in the area and the laboratory is designated a Superfund site by the Environmental Protection Agency. Laboratory officials say most of the contamination comes from aeroplane engine cleaning operations conducted during the Second World War.

Investigations at six other DoE facilities have revealed several weak points including shortages of qualified staff, undisciplined management systems and ill-defined authority and responsibility for dealing with environmental health and safety regulations.

DoE is now trying to improve its performance. A spokesman said that \$2.300 million has been appropriated in the current fiscal year to help the agency deal with environmental restoration and waste management. The proposed 1991 budget boosts this to \$2.700 million.

A draft report of the Lawrence Livermore National Laboratory findings will be prepared in early April. Other Tiger Team reviews are scheduled for the Sandia National Laboratories and the Brookhaven National Laboratory.

Robert Buder

Review needed of risks

SIR—The recent report by Gardner *et al.*¹ on the association of childhood leukaemia and employment at the British Nuclear fuels plant at Sellafield calls for an urgent examination of the current and proposed radiation protection standards. Such a review requires the data on relative risks to be converted to absolute risks so that their real significance, and the observed excess risk to be associated with a parental dose, can be assessed.

The baseline fatality risk due to leukaemia in the age range up to 25 in England and Wales can be calculated from the Office of Population Censuses and Surveys' age-specific mortality rates². In 1986, it was about 375 per million. The relative risk factors reported by Gardner *et al.*¹ can be reduced by 1 to give the excess relative risk factor which can then be applied to this baseline, given that the relative risk factor is the same for fatal and non-fatal causes.

The association with dose may be either with the dose in the few months before conception or with the total dose before conception. For the former case, Gardner's results suggest an excess risk factor of about 6.7 for a dose in excess of 10 millisieverts (a mean dose of perhaps 20 mSv) in the relevant 6 months. The excess fatality risk is then $6.7 \times 375 \times 10^{-6}$ or 2.5×10^{-8} per birth. This risk, an order of magnitude less than the risk per birth of serious congenital abnormalities, is small, but not negligible. For the second case, the excess relative risk factor seems to be about 6.3 for total doses greater than 100 mSv (a mean dose of perhaps 200 mSv), which is not significantly different. For an average paternal age of about 30 at the time of conception, the gonad dose from natural sources would be about 30 mSv. If accumulated dose is the relevant factor, it seems that paternal exposure to natural sources could well be the predominant cause of most childhood leukaemia.

The recommendations now being considered by the International Commission on Radiological Protection include occupational dose limits of 100 mSv in 5 years and 50 mSv in any year. These will ensure that almost no workers will exceed 10 mSv in the 6 months before conception or 200 mSv in total before conception. For continued occupational exposure at these limits, the commission estimates the risk of serious hereditary effects to be about 6×10^{-3} , or about 2.5×10^{-3} per child born to a highly exposed worker. The commission will need to consider whether this estimate needs to be revised in the light of the new data.

Several reservations must be recorded. The relative risk estimates have wide confidence limits. They depend on a total of 5 cases and 11 controls. Second, the relative

risk has been assessed only up to and including the age of 24. This is appropriate for leukaemia, but leaves open the possibility of a less easily observed excess risk of other cancers with longer latency periods. Third, the association of relative risk and paternal employment is also seen in workers in the iron and steel industry. A reduced relative risk is seen in coal mining. There may be causal factors other than external radiation.

H. J. DUNSTER

52 Thames Street,
St Ebbes, Oxford OX1 1SU, UK

1. Gardner, M.J. *et al.* *Br. med. J.* **300**, 423–429 (1990).
2. OPCS 1988 Mortality Statistics, Series DH2 No. 13. (HMSO, 1988).

Current opinion

SIR—In their retrospective article¹, Bynum and Heilbron's references to the electric chair are not entirely correct.

First, the company of Westinghouse took absolutely no part in the construction or installation of the electric chairs at prisons in New York State; in fact, Westinghouse did everything humanly and legally possible to block execution by alternating current^{2,3}, including seeking to abolish capital punishment in the state². The work was in the hands of an Edison associate, Harold P. Brown⁴, who secretly arranged the purchase of secondhand Westinghouse a.c. generators from Oneonta⁵ via intermediaries in Rio de Janeiro⁶. As soon as Westinghouse discovered what had happened, steps were taken to recover the generators on the grounds that they had been obtained by fraud, but the action was later dropped⁶.

Second, William Kemmler was electrocuted twice⁷ at Auburn prison, not three times, as reported by *La Nature*¹. Shortly after the first exposure to 1,400 volts for 17 seconds, Kemmler began to recover. It took a little while before the second shock, lasting over 2 minutes, was applied⁷. Edison blamed the doctors present for the fiasco; Westinghouse also attended, and commented that the doctors could have done better with an axe⁶.

The electric chair played a major role in the vicious campaign of dirty tricks waged by the Edison camp, which was anxious to promote d.c.^{1,6}, against Westinghouse and his gifted associate Tesla, who were exploiting the impressive technical advantages of a.c. Edison's purpose was to stigmatize a.c. by convincing and widely publicized demonstrations of its ability to kill^{4,8}. The medical^{8,9} and trade¹⁰ press were heavily on the side of Edison, who took detailed steps to convince the authorities of the soundness of his case^{11–14}. That almost a century later the facts of

the case should still be misunderstood is no mean testimony to Edison's talent for misrepresentation.

ROBERT JONES

30 Poplar Walk,
London SE24, UK

1. Bynum, W.F. & Heilbron, J.L. *Nature* **343**, 11–14 (1990).
2. *Lancet* **1**, 1155 (1890).
3. *Scient. Am.* **63**, 96 (1890).
4. *Scient. Am.* **61**, 240 (1889).
5. *Electr. Eng.* **15**, 296 (1890).
6. Beichman, A. *Commentary* **35**, 410–419 (1963).
7. *Med. Rec. N.Y.* **38**, 154–156 (1890).
8. White, J.W. *Med. News* **56**, 458–462 (1890).
9. MacDonald, C.F. *N.Y. Med. J.* **55**, 535–542 (1892).
10. Houston, E.J. *Electr. Eng.* **15**, 111–112 (1890).
11. *Med. Rec. N.Y.* **36**, 103 (1889).
12. Peterson, F. *Med. Record N.Y.* **36**, 223–224 (1889).
13. *Med. Rec. N.Y.* **37**, 47 (1890).
14. *Scient. Am.* **62**, 131 (1890).

Sexual attitudes

SIR—The pilot survey for the British National Survey of Sexual Attitudes and Lifestyles quoted male respondents as reporting 2.9 times as many sexual partners as women. S. J. Gurman commented (*Nature* **342**, 12; 1989) that this cannot be correct, but it is not so surprising that men have more heterosexual partners than women.

On the basis of the number of prostitutes in Switzerland¹ as a proportion of the total female population one can estimate that there are 30,946 prostitutes in Britain. In Switzerland, each has an average of one new partner every two days, so that, in Britain in one year (300 working days), there will be about 4,641,900 different sexual partnerships, equivalent to 232,095,000 partnerships during 50 years of sexually active life. This estimate implies that every British male has a lifelong average of 9 different sexual partners on the basis of prostitution alone.

If, in the survey, the number of prostitutes in the sample of women is under-represented or if the median value of partnerships rather than the average were considered, the results that give more heterosexual partners for men than for women are not so surprising, although Mrs Thatcher may doubt its value.

GIANFRANCO DOMENIGHETTI

CLAUDE FRANCE CARREL

Department of Social Affairs,
Cantonal Health Office,
Bellinzona, Switzerland

1. *Uebersicht über die Prostitution in der Schweiz* (Bundesamt für Gesundheitswesen, Bern, 1989).

Gone away . . .

SIR—With reference to J. F. Lamb's letter (*Nature* **343**, 404; 1990), I would suggest that British science is as alive and well as it has ever been. Its just that it is currently living in the United States.

M. H. GOYNS

Department of Clinical Oncology,
Royal Hallamshire Hospital,
Sheffield S10 2JF, UK

Ovchinnikov defended

SIR—We cannot accept the comments of Vera Rich (*Nature* 335, 107; 1988) about the research project on the perfluorocarbon blood substitute, Perftoran, an analogue of Fluosol-DA, Japan, undertaken at the Institute of Biophysics of the USSR Academy of Sciences. We believe that they cast undeserved suspicion on the memory of the late Professor Yuri Ovchinnikov.

In 1987, Professor Genrikh Ivanitskii, the former head of the Perftoran project, initiated in the Soviet media a campaign intended to show him as an innovator suffering for his attempts to introduce a miracle blood substitute into clinical practice, and to inspire the idea that Ovchinnikov impeded the project.

In reality, the Perftoran blood substitute suitable for use in humans never existed. Such was the conclusion of the commission of the USSR Academy of Sciences which scrupulously considered the matter and recently published a summary of the work done in *Vestnik Akademii Nauk SSR (Academy Herald, Russian)* issue 6, pp. 57–64 (1989).

Ovchinnikov, who was vice-president of the USSR Academy of Sciences responsible for biological research, did his best to provide clinicians with effective medical products. In her comments, Vera Rich simply repeated the version promulgated by Ivanitskii.

ALEXANDER A. BAEV
VADIM T. IVANOV

Shemyakin Institute of
Bioorganic Chemistry,
Academy of Sciences of the USSR,
117901 Moscow, USSR

FDA pressure?

SIR—I note with interest in the review article on “Chemical asymmetric synthesis” (*Nature* 342, 631–6; 1989) the observation that “pressure from regulatory agencies” is being exerted to “synthesize all potential new drugs as single enantiomers”. I do not know which regulatory bodies the authors have in mind, but the US Food and Drug Administration (FDA) cannot be among them.

About 25 years ago, when one of many controversies erupted about the safety of marketed drugs, I noted that one of those drugs, possessing one chiral centre, was marketed as a racemic modification. Newly arrived at the FDA, I learned early that someone at my level of reviewing chemist could never hope to persuade a pharmaceutical company to do something against its wishes, if to do so would cost it money or set back its approval. I recommended internally, therefore, that the company be

requested to resolve the compound and subject its enantiomers to tests that might show that one of the enantiomers was unnecessary or superior to the other. As a goad to accomplishing this objective, I also recommended (as a non-lawyer) that the FDA could invoke certain provisions of the Food, Drug, and Cosmetic Act, as amended, that deem a drug to be adulterated [Section 505]: “A drug . . . shall be deemed to be adulterated . . . if it has been prepared . . . whereby it may have been rendered injurious to health”; [(a) (2) (A)] and “If . . . any substance has been mixed . . . therewith so as to reduce its quality or strength [(d) (1)]”.

Although these provisions could not have been written into law with stereochemical considerations in mind, I believed them to be applicable to the question of safety of drugs with respect to the presence of an unwanted (and perhaps unnecessary) enantiomer as an adulterant.

This effort was prevented by the chief chemist of the FDA's Bureau of Drugs, who, as an inorganic chemist, could not appreciate the argument of organic chemistry, as well as being fearful of antagonizing the pharmaceutical industry. During my career in the FDA from 1963 to 1983, I know of no example of the FDA intervening to require that a company demonstrate the safety or efficacy of one enantiomer over its non-superposable mirror image or over the corresponding racemic modification. Nor have I heard subsequently of any such action.

Not only would the FDA not make any such objection, it was unable to recognize the existence of compounds with multiple chiral centres and was therefore hoodwinked into accepting at least one drug that is capable of existing as one of four diastereoisomeric pairs or mixtures thereof (see, for example, Akineton, generic name: biperiden).

BURTON L. APPLETON

PO Box 16057,
Alexandria, Virginia 22302, USA

A Pauling biography

SIR—The review by R. J. P. Williams of Serafini's Pauling biography (*Nature* 342, 135; 1989) mostly displays Williams's musings as to what makes Pauling tick. While Williams is, of course, entitled to his opinions, our image of Linus Pauling is far more positive and we are especially incensed by the biographer's view of Pauling's motivations, such as a need for tension and quarrels. We believe that Pauling was and is simply interested in everything accessible to his insights. He also fights for what he feels to be true and just. Indeed, we certainly subscribe to Williams's penultimate sentence: “Pauling . . . was for quite some time just the best of us”. And we wonder whether we have yet seen the end of his great successes.

Williams also fails to address the quality of the biography itself; it is awful. For example, periods of time are hopelessly mixed up again and again and the book is utterly slipshod otherwise (typographical errors, errors of fact and misconceptions abound: for example, column should be Coulomb and invariant [and] proportional should be inversely proportional, page 77; Ostwald should be Oswald Avery, page 152, Lorenson should be Lauritsen page 141). The book is excessively repetitive and the same interview quotations also often appear several times. We have no recollection of many occurrences Serafini relates (although one or both of us were at Caltech at the time), such as heated and loud political discussions between Pauling and Millikan in the halls of Gates and Crellin — we do not even recall ever seeing Millikan on these premises. We view as unfounded the allegations that Pauling was unsympathetic towards organic chemistry or the humanities division. And so on. So, as to both the facts and to the flavour of things in the Pauling days, we fear that Serafini's biography is totally unreliable.

JÜRIG WASER

6120 Terryhill Drive,
La Jolla, California 92037, USA

VERNER SCHOMAKER

California Institute of Technology,
Pasadena, California 91125, USA

Fetus research

SIR—In a recent Commentary (*Nature* 342, 469; 1989), reference was made to the report of the Polkinghorne Committee on the use of fetuses and fetal material for research. Research on living intact aborted fetuses is apparently forbidden, but the reported criteria for death are absence of heartbeat or of spontaneous respiration. (The ‘or’ may be significant). Fetal heartbeat may be difficult to detect, and capability for spontaneous respiration is rarely established until the seventh month of gestation, the legal limit for abortion in the United Kingdom anyway. Experimentation on a spontaneously breathing fetus would then in law constitute infanticide.

Far from presenting ‘strict guidelines’, the Polkinghorne criteria are therefore the minimal precautions to be observed by an experimenter to avoid not just controversy but also the possibility of criminal prosecution. Had the concern been to avoid the possibility of distress to a sentient experimental subject, the appropriate guideline would be the definitive and confirmed cessation of activity in the central nervous system, as is the general requirement on human subjects.

A. J. McEvoy

Institut de Chimie Physique,
EPFL-Ecublens,
CH-1015 Lausanne, Switzerland

Hands up for the Gaia hypothesis

James E. Lovelock

The concept of Gaia, a self-regulating Earth, excites both admiration and obloquy. Its inventor (or rather re-discoverer) describes the genesis and evolution of the hypothesis.

In essence, the Gaia hypothesis was born 25 years ago. Those who have since worked closely with it see it as reasonable science and growing ever more plausible as evidence and models come together. But theirs is a minority opinion. Most scientists regard Gaia as little more than metaphor; some even denounce it as anti-science. This Commentary is about the history, the arguments, the evidence and the criticisms of Gaia, and about where it now stands.

Beginnings

The history of the idea goes back to the mid-1960s, to the Jet Propulsion Laboratory in California. The National Aeronautics and Space Administration had sought my help in the quest to discover whether there was life on Mars, and in 1965 I proposed some physical tests for the presence of planetary life¹. One of these was a top-down view of the whole planet instead of a local search at the site of landing. The test was simply to analyse the chemical composition of the planet's atmosphere. If the planet were lifeless, then it would be expected to have an atmosphere determined by physics and chemistry alone, and be close to the chemical equilibrium state. But if the planet bore life, organisms at the surface would be obliged to use the atmosphere as a source of raw materials and as a depository for wastes. Such a use of the atmosphere would change its chemical composition. It would depart from equilibrium in a way that would show the presence of life.

At that time Dian Hitchcock joined me, and together we examined atmospheric evidence from the infrared astronomy of Mars². We compared this evidence with that available about the sources and sinks of the gases in the atmosphere of the one planet we knew bore life, Earth. We found an astonishing difference. The atmosphere of Mars was close to chemical equilibrium and dominated by carbon dioxide, but that of Earth was in a state of deep chemical disequilibrium. In our atmosphere carbon dioxide is a mere trace gas. The coexistence of abundant oxygen with methane and other reactive gases would be impossible on a lifeless planet. Even Earth's abundant nitrogen and water are difficult to explain by geochemistry. There are no such anomalies in the martian and venusian atmospheres, and their existence in the Earth's atmosphere signals the presence of living

organisms at the surface. Sadly, we concluded, Mars is probably lifeless.

That first sight from space of a dappled white-and-blue sphere irreversibly altered most people's emotions about the Earth. Atmospheric chemistry, seen from above,

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

James Lovelock — a top-down view.

was in scientific terms for me a similar revelation about the Earth. This top-down analysis revealed the atmosphere as a gas mixture like that of the intake manifold of an internal combustion engine, with oxygen and combustible gases mixed, and very different from the exhausted, carbon-dioxide-dominated atmospheres of Mars and Venus. Further, I knew that the chemical composition of the atmosphere was stable for long periods when compared with the residence times of its gases. One afternoon in 1965, thinking about these facts, the thought came that such constancy requires the existence of an active control system.

At that time I lacked any idea of the nature of this control system, except that the organisms on the Earth's surface were part of it. I learned from astrophysicists that stars increase their heat output as they age and that our Sun has grown in luminosity by 25 per cent since life began. I realized that, in the long term, climate also might be actively regulated. The notion of a control system involving the

whole planet and the life upon it was now firmly established in my mind. Sometime near the end of the 1960s I discussed this idea with my near neighbour, the novelist William Golding. He suggested the name Gaia as the only one appropriate for so powerful an entity. Here I shall use the word Gaia as shorthand for the theory.

I first formally put forward the idea of Gaia as a control system in 1972³. Shortly afterwards I began to work with Lynn Margulis, a collaboration that has continued. We first stated our hypothesis in words such as "Life, or the biosphere, regulates or maintains the climate and the atmospheric composition at an optimum for itself". By 1973 we refined the hypothesis and re-stated Gaia in papers in *Tellus*⁴ and *Icarus*⁵ as "The notion of the biosphere as an adaptive control system that can maintain the Earth in homeostasis, we are calling the Gaia hypothesis".

Teleology

Neither Lynn Margulis nor I have ever proposed a teleological hypothesis. Nowhere in our writings do we express the idea that planetary self-regulation is purposeful, or involves foresight or planning by the biota. It is true that our early statements about Gaia were imprecise and open to misinterpretation, but this does not justify the persistent, almost dogmatic, criticism that our hypothesis is teleological. Sometimes the criticism reveals an ignorance of the circular logic of control theory — how often are the terms negative and positive feedback used, through lack of understanding, as mere metaphors?

As an inventor I know it is easy to image, construct and then reduce a self-regulating device to practice. I also know how difficult it is to explain it in words or even in analytical mathematics. In many ways Gaia, like an invention, is difficult to describe. The nearest I can reach is to call Gaia the theory of an evolving system — a system made from the living organisms of the Earth, and from their material environment, the two parts being tightly coupled and indivisible. This evolutionary theory views the self-regulation of climate and chemical composition as emergent properties of the system. The emergence is entirely automatic; no teleology is invoked. Gaia evolves as a system, gradually, during long periods of homeostasis that are punctuated by sudden simultaneous changes in both organisms

Sandy Orchard

and environment. Such changes move the system to new and different homeostatic states; a significant jump of this kind occurred between the anaerobic Archaean and the oxygenated Proterozoic 2.5 thousand million years ago. Gaia offers a resolution of the long debate over whether evolution was gradual or punctuated by suggesting that it was both.

This is not a new theory of evolution. James Hutton⁶ intuited it when, in 1788, he saw the Earth as a super-organism which could only be studied properly by physiology. Alfred Lotka⁷ also expressed it in 1925 when he clearly stated that the evolution of organisms could not be separated from the evolution of their physical environment. Gaia theory is not incompatible with Darwin's great vision; it includes the evolution of the organisms by natural selection as an essential part of a self-regulating planet.

Much of the confusion about Gaia and darwinism comes from the misuse of the concept of adaptation. My friend and colleague Andrew Watson succinctly expressed the step that distinguishes Gaia from darwinism in a debate before the Linnean Society in December 1989. It lies in the tightness of the coupling between the organisms and their physical environment. Watson observed that almost everyone now accepts that life profoundly influences the environment; this is now the conventional wisdom among geochemists, and a considerable change from their view pre-Gaia. It is equally obvious that life is influenced by and adapts to the environment. This is the older wisdom that has prevailed throughout this century. Therefore, as Watson pointed out, life and the environment are a coupled feedback system, where changes in one element will affect the other and this may in turn feed back on the original change. The real debate, then, is how important and how tight is the coupling? Does it, as we believe, confer new properties on the system, such as enhanced stability or behaviour similar to that of a living organism?

Hollism

Gaia asserts that this close coupling of organisms and their environment is strong enough to have greatly influenced the way in which the life-environment system on Earth, and on other planets with life, has evolved. It is strong enough that we will not properly understand Earth history until we think of the system as just that, a whole system, and stop trying to understand its parts in isolation from one another.

Like living organisms generally, Gaia is an open but bounded system. On one side, the bound is space and Gaia exchanges radiation across it. The other bound is the base of the crustal rocks where matter exchanges with the near-infinite mass of

Earth's mantle. Gaia's wastes are low-quantum-energy infrared radiation that escapes to space, and rock subducting to merge with the mantle. The profound disequilibrium of the atmosphere is one measure of the entropy reduction of this system. I think that James Hutton was right to call the Earth a super-organism and to regard physiology as the proper science for its investigation.

In 1982, Richard Dawkins⁸ made the strong criticism that no way exists for evolution by natural selection to lead to altruism on a global scale. My answer was the numerical model, daisyworld^{9,10}. The model is of a simple planet that keeps its climate constant in the face of ever-increasing solar output. It does so by the competition for space of two species of daisies, one dark and one light. The planetary albedo is tightly coupled to the evolution of the daisies, and the evolution of the daisies is tightly coupled to the evolution of the climate.

More sophisticated models have been developed that go beyond daisyworld and illustrate the simultaneous self-regulation of climate and chemistry by bacterial ecosystems. These new models of Gaia are general and not limited as are models drawn from the separated disciplines of science. They can include many species or ecosystems simultaneously and follow their regulation as well as that of the material environment. Population biologists, for 60 years, have been unable to model more than two species simultaneously. The box models of biogeochemists are also unstable and unduly sensitive to the choice of initial conditions. These modellers handicap themselves by failing to include the biota, or the material environment, interactively.

Some geochemists^{11,12} have been opponents of Gaia from the beginning. Their criticism is the reasonable one that Gaia is not needed to explain the facts of the Earth. Yet these scientists are ready to tolerate Gaia as a theory under test and interesting as a source of ideas for new experiments. In an act of courage and generosity, the climatologist Stephen Schneider organized a conference of the American Geophysical Union in 1988 solely to discuss Gaia. The conference is described in Schneider's book *Global Warming*¹³. This meeting marked the end of the false accusation of teleology — what some, illogically, have called the strong Gaia hypothesis. At the conference, J.W. Kirchner made a spirited attempt to demolish all notions of Gaia. Like some figure of the Inquisition, he publicly burned several imaginary Gaia's, and his pyrotechnic demolition of the strong Gaia stole the show. But when the sparks faded, the real system Gaia was still there hidden only by the smoke. The flux of papers inspired by Gaia, and now appearing in the journals, are the real proof of

the value of the conference. It has not stopped peer review from censoring any mention of Gaia by name.

The value of Gaia as a theory is illustrated by the interpretation of the long-term history of climate it makes possible. The conventional wisdom about the problem of the cool early Sun is that a warmth sufficient for the start of life, some 3.6–4 thousand million years ago, came from a greenhouse blanket of atmospheric CO₂. This accounts for the start, but there has been no satisfactory explanation of the constancy of the Earth's climate throughout history.

Geophysiology

Geophysiology — my own neologism, but a useful one I think — can provide such an explanation. In a model of the Archaean period^{14,15}, a symbiotic pair of ecosystems, photosynthesizers and anaerobes, evolve tightly coupled to their physical and chemical environment. The model system sustains a climate and atmospheric composition comfortable for Archaean life. The photosynthesizers tend to cool by removing CO₂, both directly and by increasing the rate of weathering. The anaerobes tend to warm by producing the greenhouse gas methane. By including a third ecosystem (consumers), the Archaean model is extended into the Proterozoic without loss of self regulation. These models are stable, indifferent to the choice of initial conditions and can resist perturbation.

Among the insights that come from a gaian approach are that planetary life can never be sparse. A planet with sparse life could never self regulate. The geophysical and geochemical evolution of the terrestrial planets is progressive and towards states like those of Mars and Venus now. During this evolution there will be a period when conditions are favourable for life and, to survive, the organisms must become abundant enough to affect and become coupled to the geochemical evolution; mere adaptation is not enough. If they fail, planetary conditions will continue to change inorganically until the point is reached when life is impossible.

Another insight from Gaia is that the biota are continuously engaged in the removal of CO₂ from the air and that the present cool climate results from such biological pumping¹⁶; viewed in this way, the warmth of the present interglacial perhaps implies a less-abundant or less-efficient biota. Gaia theory also led directly to the identification of another possible climate-control system — the global-scale emission of dimethyl sulphide from the oceans, which was discovered in the course of a search for a chemical agent of biological origin to complete the natural sulphur cycle¹⁷. This process could turn out to be as great a climate regulator as the carbon dioxide and methane greenhouse.

Some predictions from Gaia

Prediction and year	Test and result
Mars lifeless, from atmospheric evidence (1968)	Viking mission (1977) — strong confirmation (ref. 2)
Organisms make compounds that can transfer essential elements from ocean to land surfaces (1971)	Dimethyl sulphide and methyl iodide both found (1973; ref. 17)
Climate regulation by control of CO ₂ through biologically enhanced rock weathering (1981)	Microorganisms greatly increase rock weathering (1989; refs 16, 20)
Climate regulation through cloud density control linked to algal sulphur gas emissions (1987)	Still under test (ref. 21)
Oxygen has stayed at 21 ± 5 per cent for past 200 million years (1973)	New hypotheses for regulation by fires, phosphorus cycling (refs 14, 22)
Archaean atmospheric chemistry dominated by methane (1988)	Still under test (refs 10, 14)

From a Northern Hemispheric human view we see glaciations as a disaster, but from a planetary viewpoint they may be a preferred state, one where life is more vigorous. When seen in this light, the current interglacial appears as a planetary fever, a pathology. In this context, pollution with greenhouse gases and the widespread destruction of natural habitats are further insults to a weakened system. Broecker¹⁸ has warned that geophysical processes alone could lead to "surprises in the greenhouse". As an active control system, Gaia can speed and intensify such surprises. A systems response might at first mask the expected warming by the greenhouse, for example by an actively increased cloudiness. We might then become complacent, or argue that the greenhouse has been overstated, only to be surprised by catastrophic climate change when the system is overwhelmed.

Testability

Gaia theory is testable, and the tests carried out so far are listed in the table. Why then has Gaia been so unpopular among scientists? An obvious reason would be the natural inertia of science which means that large-scale theories are digested only slowly. It took 40 years for another Earth theory, plate tectonics, to be accepted. With Gaia there are, I think, additional reasons for its unpopularity.

Gaia is developing normally in the Earth sciences and in time will be either accepted or rejected on the evidence. Some biologists by contrast denounced the idea even before evidence was available. Postgate¹⁹ referred to Gaia as "silly and dangerous. . . pseudoscientific myth making". He went on to blame Gaia for the disrepute now attending science.

Even my friends among biologists become anxious when I talk of Gaia as an organism. Yet in so doing I merely continue a tradition that goes back at least as far as Leonardo da Vinci and that was made scientific by James Hutton. Biologists rejected this tradition in the nineteenth century when their subject first

became intensely reductionist. Biology has seemed to be imprisoned in a narrow, almost puritan, reductionism, a consequence perhaps of the long and continuing war with vitalists, animists and religious fundamentalists. In war, it is said, the first casualty is truth. I think that science, when it engages in adversarial encounters with crude dogma or non-science, loses objectivity and itself grows dogmatic. In their criticisms, biologists seem to see Gaia as another threat to the sanctity of their creature, neo-darwinism. No one denies the power of reduction with its apotheosis in the triumphs of molecular biology. But the reductionists in their dogmatic crusade have made holism a pejorative term, and now cybernetics, Gaia and systems analysis are categorized with 'new age' philosophy and organic food. No wonder the greens prefer Gaia to biology.

A great loss results from this rejection of a rigorous top-down approach to systems. Holism is complementary not adversarial to reduction. What engineer now would dare to dissect a working computer into its hardware components before testing it as a whole system? Without holistic thinking, physiology, engineering and Gaia would be completely disabled.

The most difficult and the dirtiest criticism of Gaia is that there is nothing new in the idea, that it has all been said before — difficult, because there is much truth in it; dirty, because, as William James said of the fate of any new idea, "First, it's absurd, then maybe, and last, we have known it all along". Most biologists and many geochemists are ignorant of the details of control theory, the domain of engineers and physiologists. Merely for them to say it was explained by the balance of nature long ago is not enough. If it were, then explanations of why the concentration of oxygen is at 21 per cent, or of why the climate has remained favourable for 3.6 thousand million years, would already be available. Before the advent of the Gaia hypothesis, such questions were rarely asked, and

would have been as pointless as asking an anatomist or a biochemist how human temperature is regulated. Such questions about systems cannot be answered from the separated disciplines of biochemistry or biogeochemistry, nor from neo-darwinist biology. The answer comes from physiology or control theory.

What of the criticism that Gaia gives support to anti-science? True, Gaia as a name, or a sign, has extended far beyond my intentions. As the semiotician Myrdene Anderson put it several years ago, "Gaia is an empty sign with near infinite capacity for signification". I watch it filling fast, and mostly with rubbish, like an empty skip left on a London street. Surely, though, this is the fate of any new sign and has nothing to do with the quality of the science of Gaia. I have offered 'geophysiology' as an alternative name, but so far there have been few takers.

I am not dismayed by the wide general interest in Gaia; rather, I am moved by the interest shown by theologians and philosophers. Whether or not Gaia is an accurate description of the Earth, it forces a different view from that of conventional wisdom — a view that could be crucial to understanding the consequences of pollution and environmental disturbance.

Arthur Clarke's notable comment, "How inappropriate to call this planet Earth when clearly it is Ocean", illustrates the wisdom of a top-down view. Few have been privileged, as astronauts, to see the Earth in its splendour from above. But anyone can rise above dogma, scientific or religious, and look down to see and cherish a most seemly Earth. □

James E. Lovelock is a Visiting Professor in the Department of Cybernetics, University of Reading, Whiteknights, Reading RG6 2AL, UK.

- Lovelock, J.E. *Nature* **207**, 568–570 (1965).
- Hitchcock, D.R. & Lovelock, J.E. *Icarus* **7**, 149–159 (1967).
- Lovelock, J.E. *Atmos. Environ.* **6**, 579–580 (1972).
- Lovelock, J.E. & Margulis, L. *Tellus* **XXVI**, 1–10 (1974).
- Margulis, L. & Lovelock, J.E. *Icarus* **21**, 471–489 (1974).
- Hutton, J. *Trans. R. Soc. Edinburgh* **1**, 209–304 (1788).
- Lotka, A.J. *Elements of Physical Biology* (Williams & Wilkins, Baltimore, 1925).
- Dawkins, R. *The Extended Phenotype* (Oxford University Press, 1982).
- Lovelock, J.E. in *Biological Mineralisation and Biological Metal Accumulation* (eds Westbroek, P. & de Jong, E.W.) 15–25 (Reidel, Dordrecht, 1983).
- Watson, A.J. & Lovelock, J.E. *Tellus* **35B**, 284–289 (1983).
- Holland, H.D. *The Chemical Evolution of the Atmosphere and the Oceans* (Princeton University Press, 1984).
- Walker, J.C.G., Hays, P.B. & Kasting, J.J. *geophys. Res.* **86**, 9776–9782 (1982).
- Schneider, S.H. *Global Warming* (Sierra Club Books, San Francisco, 1989).
- Lovelock, J.E. *The Ages of Gaia* (W.W. Norton, New York, 1988).
- Lovelock, J.E. *Rev. Geophys.* **27**, 215–222 (1989).
- Lovelock, J.E. & Whitfield, M. *Nature* **296**, 561–563 (1982).
- Lovelock, J.E., Maggs, R. & Rasmussen, R.A. *Nature* **237**, 452–453 (1972).
- Broecker, W.S. *Nature* **328**, 123–126 (1987).
- Postgate, J. *New Scientist* **60** (7 April, 1988).
- Schwartzman, D.W. & Volk, T. *Nature* **340**, 457–459 (1989).
- Carlson, R.J., Lovelock, J.E., Andreae, M.O. & Warren, S.J. *Nature* **326**, 655–661 (1987).
- Kump, L. *Nature* **335**, 152–154 (1988).

Engineer in the White House

The Chief of Staff at the White House is a powerful figure. The present incumbent, Mr John Sununu, has the distinction of having himself run for elected office and of being an engineer.

Washington

MR John Sununu, Chief of Staff at President George Bush's White House, is probably the first person with a technical background in such an influential position since Benjamin Franklin officiated at the court of George Washington.

Sununu is an engineer who taught at Tufts University and ran a consulting firm before being elected governor of New Hampshire in 1982. He is pleased with his foray into elected office. Before running for the governorship of New Hampshire, he had served for only about a year as a member of the state legislature. He seems particularly pleased that he was able to defeat an incumbent to become chief executive of the state which is fiftieth (out of 52) in size.

But now, at least at first blush, Sununu is formally a 'facilitator' — Washington's word for those who make things happen, supposedly without much regard for what they are. Operationally, he is the president's door-keeper, but he claims to differ from his predecessors in not trying to keep people out, but to gather them in instead. One of the services a technical person can uniquely provide for a president, Sununu says, is to demonstrate that real-world problems are complex and usually threaded with uncertainty. Getting the president to meet people with conflicting views is part of the job.

Yet there is more than that to the job. Sununu puts it modestly: "Sometimes the president asks for my advice, and then I give it". That seems to happen often. In the past few weeks, as the administration's chief link-man with the Congress on the Clean Air Bill, inevitably seized upon by faddish interests as a vehicle for special-interest amendments, but which also includes regulations on acid rain and motor-car emissions, Sununu has earned the reputation of dragging the administration's feet on motor-car emissions and for shrinking from proposals that the United States should unilaterally cut back on the emission of greenhouse gases.

It is relevant that Sununu appears to be a true conservative. On the parlous state of public education in the United States, for example, he agrees that there is a problem, but insists that there is only one solution, which lies beyond the sphere of action of the 'feds' (the federal authorities). Instead, he offers the school systems of New Hampshire, now "Number one in SAT [college entrance examination] scores", as models of good practice and,

incidentally, as proof of his own attainments.

The trick is to create the circumstances in which there is a tight relationship between parents, teachers and children, with a natural feedback between the interests of the three parties so that the students learn to learn. "It's not a question of money", he says, "but of attitudes." State and local government can provide only "stimulus for involvement". "I'm convinced I can teach any child anything if he wants to learn." Sununu did not say, and was not asked, whether he knew that this is the principle on which the British government's Education Reform Bill, as it applies to schools, is founded.

But what about exhortation, which is not denied the federal government by states' rights? Like other administration officials these days, Sununu enthuses about last autumn's governors' conference at which Bush spent two days talking at and with state governors about the importance of public education. He says it mattered most to those in attendance that the president had judged the occasion worth two days of his time.

Whether there has yet been time for that warm feeling to have filtered through to the depressed urban school systems is doubtful. There is no prospect of a national curriculum (as there is in Britain), but Sununu sets great store by a system now being devised for monitoring the performance of school systems centrally, so that we can "know where we are".

Sununu's "feds off" philosophy applies in other fields. He has stoutly resisted, on behalf of the administration, appeals for special assistance for particular industries (semiconductors, for example) and industrial projects (the development of high-definition television). But he supports basic research and even a modest spillover into development spending.

Environmental issues are more complicated, and Sununu's record as governor of New Hampshire is not such as to endear him unambiguously to the Sierra Club or the Audubon Society. He successfully carried through legislation for the abatement of acid rain, but was also a supporter of the Seabrook nuclear plant (which has just been awarded an operating licence, see page 96).

More recently, several environmental groups have been angered by reports that Sununu was responsible for the lukewarm

commendation of the need for action to avoid climatic change in Bush's speech last month to the International Panel on Climate Change (IPCC) meeting in Washington.

Sununu is unrepentantly forthright, saying that the president's appearance at the IPCC was a sign of the administration's serious regard for the possibility of climatic change, as is its spending of more than \$1,000 million on relevant research. But with the federal government's obligations to people's jobs and prosperity, it would be wrong to distort the present economic pattern without firm evidence that climate change will actually be induced by the greenhouse gases.

Neither Sununu nor the administration is opposed to negotiations on an international convention — indeed, they support the idea — but the modellers have not yet convinced him that greenhouse warming will come to pass. "I'll be convinced when they can put in the models realistic climate data for 100 years ago, and then run them to reconstruct the climate history we know." With some glee, he fishes in a pile of papers for a press release reporting a decline of land and sea-surface temperature in the south-eastern United States during the past half century, "just where they say it should be warming".

Sununu sees his role in this internal debate as typifying the proper role of technical people in government. Politicians are mostly lawyers, he says, and lawyers are used to selecting facts to bolster up the case they have to argue. But technical people turn the process upside-down, and argue a case only when they have the facts. "Politicians need help", he says. So why are there a good many scientists in the field apparently persuaded that the greenhouse effect is real? "Well", after a little thought, "every profession has its zealots."

Whatever his profession, and despite his reputation, Sununu seems not to be one of these. He is a combative man in a powerful post, and no doubt can be an infuriating opponent, but he seems to bear the cares of office without serious damage to his good humour.

The risk he runs is that, with the president's popularity rising steadily in the polls, the public will need somebody else to blame for departures from expectations of the ideal. One who talks of the "pressures of reality" may be typecast for that role.

John Maddox

Greenhouse cooling up high

Ralph J. Cicerone

CHANGES in the atmosphere's chemical composition are stimulating a lot of careful thought: how far will thinning of the ozone layer in the stratosphere proceed?; will radiative forcing, the growing greenhouse effect, accelerate in the troposphere? The consensus is that it is at least plausible that continued increases, due to human activity, in concentrations of carbon dioxide, methane and chlorofluorocarbons, will lead to measurable, if not damaging, environmental disturbances. Roble and Dickinson¹ now widen the debate by considering the effect of emissions of carbon dioxide and methane on the upper atmosphere. Dramatic consequences could ensue: the upper reaches may cool by tens of degrees and pressures will be reduced; concentrations of oxygen atoms will be reduced and those of hydrogen atoms increased.

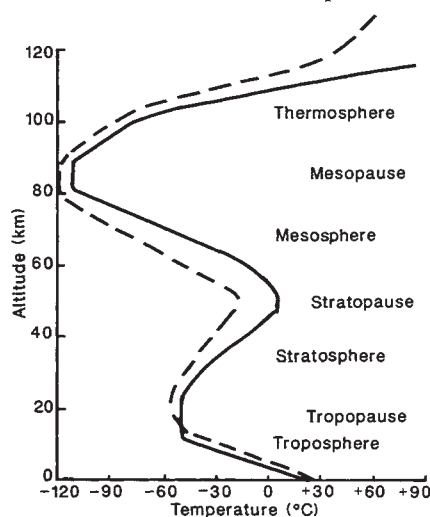
This is not a wild proposition, but rather a simple and direct one. The increased concentration of CO₂ in the stratosphere causes cooling because CO₂ is an efficient radiator to space of energy through infrared emissions (the obverse side of the greenhouse coin in which infrared activity causes tropospheric warming through radiation trapping). This is thought to be the surest consequence² of increased CO₂; this cooling should extend into the lower mesosphere³. The concentration of hydrogen atoms in the thermosphere is known to be sensitive to the amounts of water vapour and methane at the top of the stratosphere⁴.

Also, the link between the ground-level concentration of CH₄ and the stratospheric concentration of water vapour has been known for a while. Yet no one had looked at the very high atmosphere to see how its chemistry and physics might be affected by ground-level changes in CO₂ and CH₄. At times like these we see how fragmented are our scientific communities and our knowledge of the chemistry of the Earth's atmosphere.

The upper atmosphere is an arena for a rich variety of phenomena, much studied, such as visible aurorae over the polar regions, airglow optical emissions and electrical current flows due to charged particles. Commercial and military communication through radio waves reflected by the ionosphere, starting at the mesosphere, added to the scientific interest in the upper atmosphere. Intensive scientific research has focused on the atmosphere above about 100 km altitude. Instruments flown by satellites have probed the ionosphere and the upper atmosphere with remote sensing by ultraviolet and visible light and *in situ* mass spectrometers, as have ground-based

radars and optical airglow sensors. Its gross morphology (temperature, degree of ionization and chemical composition) and variations have been mapped. Atomic and molecular atmospheric processes such as those involved in photoionization of gases and recombination of ions have been studied in the laboratory, and macroscopic collective phenomena have been modelled and observed, such as gravity waves, plasma waves and electrical currents.

Numerical models of the ionosphere and upper atmosphere have progressed dramatically. Fluid dynamical, thermal balance and chemical kinetic equations are solved for winds and temperatures,



Variation of temperature with altitude in the atmosphere (solid line) revealing the various strata. Precise details vary with latitude and season. The temperature of the thermosphere rises to over 700 °C at 250 km. Above the mesopause, solar radiation ionizes atmospheric gases to create the ionosphere, divided into sub-layers D, at 90 km, E, at 110 km, F, above 200 km, defined by differing behaviour. Roble and Dickinson are the first to study the effect of CH₄ and CO₂ on the upper reaches of the atmosphere. The broken line shows the predicted effect of increased CO₂, which extends to cooling by 50°C at the top of the thermosphere.

given fairly basic input data for solar input and the major chemical species in air. Roble and Dickinson and their colleague, E.C. Ridley, and others can now calculate reasonably realistic dynamical patterns, ion types and numbers and temperature variations, for example, from solar-activity maxima to minima. Although some boundary conditions for the equations are still specified too rigidly, many of the phenomena are simulated amazingly well and the coupling of fluid dynamics, radiation, charged-particle dynamics and chemical reactions in

modern models is becoming very realistic.

Roble and Dickinson have introduced into such a model changes from the reference-state concentrations of CO₂ and CH₄ of the 1950s; they doubled and halved the observed values at the model's lower boundary (62 km). Higher concentrations for these gases cause the computed temperatures to fall by 50 K near 400 km altitude as compared to the reference state, whereas lower concentrations cause temperatures there to rise by 35 K. Concentrations of the dominant gas, oxygen atoms, fall by almost 40 per cent near 400 km when CO₂ and CH₄ are doubled and they rise by the same amount when the driving gases are halved. Hydrogen atom concentrations either rise by 50 per cent or fall by 20 per cent for the same changes.

The potential consequences of such perturbations are numerous. Roble and Dickinson mention some of them for the case of increasing CO₂ and CH₄: atmospheric drag on orbiting satellites will be reduced as the density of the upper atmosphere decreases and noctilucent clouds in the mesosphere will become more common as mesospheric water vapour quantities increase and temperatures decrease simultaneously. Ionospheric ion densities would be reduced near the E- and F-region peaks, and topside plasma scale heights will shrink. Global circulation and spatial morphology of the thermosphere may also change.

All of these and other perturbations must be analysed more thoroughly; Roble and Dickinson have made only the beginning. Various interactions in the system could lead to altered predictions; effects could be larger or smaller. For example, wave propagation and dynamical heating through the mesosphere and thermosphere could be altered as vertical temperature profiles change. One notable inconsistency in the paper leads to underestimated effects. The authors held the water-vapour mixing ratio fixed at 6 parts per million (p.p.m.) at 62 km; it should have risen to 9 p.p.m. when methane was doubled. Thus, hydrogen-containing species should increase in concentration even more than the authors predicted and hydrogen escape rates will increase (at least that part of the escape flux that is thermal^{5,6}) and the ratio of hydrogen/deuterium escape rates will change⁷. Further, these larger amounts of mesospheric H₂O and simultaneously lower temperatures will increase the regularity of noctilucent clouds⁸ and possibly alter D-region ion chemistry. It is also conceivable that the added water could lead to more cloud formation in the polar stratosphere, thus priming the system for ozone destruction⁹, but the size of this effect remains to be determined. In the thermosphere, as the H/O ratios must change by perhaps a factor of three, the redox state

of the high atmosphere would tend to reducing rather than oxidizing. Results could depend also on the phase of solar activity.

The driving perturbations and the proposed effects do deserve more study. CO₂ and CH₄ should be simulated separately to isolate mechanisms of change, the sizes of the changes in temperature, pressure and mixing ratios of chemicals should be compared to corresponding ranges of natural variability so that trends can be extracted from observations, and inconsistencies in the calculations should be removed (mesospheric concentrations of CO₂ were closer to 300 than to 330 p.p.m. in the 1950s). Also, further attention should be paid to low CO₂ and CH₄ amounts; concentrations of these species have been as low as 190–280 p.p.m. (ref. 10) and 0.35 p.p.m. (ref. 11) respectively in the past 160,000 years and well outside this range for CO₂ at earlier times.

Roble and Dickinson's report opens our eyes to further disturbances to the global atmosphere that are primarily due

to human activities. These may turn out to be manageable or even insignificant. But like the ozone hole over Antarctica, they may be large enough to be seen from Mars or some other vantage point in space. □

Ralph J. Cicerone is in the Geosciences Department, University of California, Irvine, California 92717, USA.

1. Roble, R.G. & Dickinson, R.E. *Geophys. Res. Lett.* **16**, 1441–1444 (1989).
2. Mahlman, J.D. *Testimony to US Senate Committee on Commerce, Science and Transportation, Subcommittee on Science, Technology and Space* 8 May 1989, 122–139 (US Govt Printing Office, 1989).
3. Brasseur, G.P. & Hitchman, M.H., *Science* **240**, 634–637 (1988).
4. Ehhalt, D.H. *J. geophys. Res.* **91**, 2843 (1986).
5. Hunten, D.M. & Strobel, D.F. *J. Atmos. Sci.* **31**, 305–317 (1974).
6. Liu, S.C. & Donahue, T.M. *J. Atmos. Sci.* **31**, 2238–2242 (1974).
7. Yung, Y.L. *et al.* *J. geophys. Res.* **94**, 14971 (1989).
8. Thomas, G.E., Olivero, J.J., Jensen, E.J., Schroeder, W. & Toon, O.B., *Nature* **338**, 490–492 (1989).
9. Blake, D.R. & Rowland, F.S. *Science* **239**, 1129 (1988).
10. Barnola, J.M., Raynaud, D., Korotkevich, Y.S. & Lorius, C. *Nature* **329**, 408–414 (1987).
11. Chappelaz, J., Barnola, J.M., Raynaud, D., Korotkevich, Y.S. & Lorius, C. *Nature* (in the press).

GENE TARGETING

Tapping the cellular telephone

Mario Capecchi

GENE targeting means that we now have the potential to generate mice of virtually any desired genotype. In the first instance, standard recombinant DNA technology is used to alter a cloned DNA sequence of a chosen locus; the modified DNA is then introduced into a pluripotent stem cell derived from a mouse embryo, and homologous recombination between the exogenous and endogenous chromosomal sequence transfers the mutation to the genome. Microinjection of the stem cells containing the modified locus into mouse blastocysts is used to generate germ-line chimaeras. Finally, interbreeding of heterozygous siblings yields animals homozygous for the desired mutation. This technology has already been used to create germ-line chimaeras containing targeted disruptions in the *HPRT*, *abl*, *en-2*, *n-myc*, β -2 microglobulin, *igf-2* and *int-1* genes. Despite these achievements, we know very little about the recombination mechanisms underlying gene targeting in mammalian cells. On page 170 of this issue¹, Zheng and Wilson provide an elegant demonstration of an unexpected feature not predicted by any simple recombination model, namely that the gene targeting frequency in mammalian cells is independent of the number of target sequences present in the genome. The authors report that the efficiency of targeting into the dihydrofolate reductase gene (*DHFR*) was identical in a normal and in an amplified Chinese hamster-

ovary cell line containing 400 copies of the *DHFR* gene. Each *DHFR* gene in the amplified cell line is presumed to be equivalent as the *DHFR* enzyme and messenger RNA levels are proportional to the number of modified genes.

This observation is unexpected because the exogenous DNA sequence must search an enormous number of DNA sequences to find the cognate chromosomal sequence and participate in homologous recombination. In addition, the gene targeting frequency is independent of the number of copies per cell of exogenous DNA molecules introduced into the recipient cell². This means that the search for the cognate chromosomal DNA sequence is probably not rate-limiting. Furthermore, because this gene targeting occurs at readily detectable frequencies, there is probably a cellular machinery for the efficient sampling of chromosomal sequences.

Interestingly, as the authors point out, in yeast the gene targeting frequency seems to be proportional to the number of target sequences in the genome^{3,4}. For example, targeting into the ribosomal RNA genes, which are present at 140 copies per genome, is 100–200 times more frequent than into the *leu-2* gene. Gene targeting in yeast and mammalian cells also differs in a number of other respects. The frequency depends more on the extent of homology between the exogenous and chromosomal sequences in mammalian cells than it does in yeast^{5–7}:

in yeast this dependence is linear, but in mammalian cells it is exponential. On the other hand, the presence of homologous ends in the targeting vectors is apparently more critical for gene targeting in yeast than in mammalian cells^{8,9}. But in spite of these differences, the absolute frequencies of gene targeting in yeast and mammalian cells are surprisingly comparable. The main difference between the two cell types is that when they are exposed to exogenous DNA, yeast mediates recombination events that are almost exclusively homologous, whereas recombination events in mammalian cells are predominantly non-homologous. So the problem in mammalian cells is to identify homologous recombination in a vast arena of scattered, non-homologous recombination activity.

In any organism, gene targeting exploits the cellular machinery to mediate homologous recombination between exogenous and endogenous DNA. What is the normal function of this machinery? One function must be DNA repair, but evidence is accumulating that all the DNA sequences in a cell may be continually in communication with one another. In yeast, gene conversion occurs between cognate DNA sequences on homologous or non-homologous chromosomes with comparable frequency¹⁰. In humans, examples are turning up in which the normal gene seems to have been mutated by gene conversion as a result of a pseudogene residing on non-homologous chromosomes. If all DNA sequences within a cell are continually talking to each other, then the function of the communication machinery could extend well beyond DNA repair. For example, it might participate in the generation, maintenance and divergence of gene families and in the shuffling of exons between genes sharing stretches of homologous DNA sequences.

At present there is a flurry of activity to employ gene targeting in pluripotent stem cells as a means to define the function of a wide range of genes in the living mouse. Our increasing understanding of the gene targeting reaction itself will continue to improve the design of such experiments. □

Mario Capecchi is at the Howard Hughes Medical Institute, University of Utah, 337 South Biology, Salt Lake City, Utah 84112, USA.

1. Zheng, H. & Wilson, J.H. *Nature* **344**, 170–173 (1990).
2. Thomas, K.R., Folger, K.R. & Capecchi, M.R. *Cell* **44**, 419–428 (1986).
3. Petes, T.D. *Cell* **19**, 765–744 (1980).
4. Szostak, J.W. & Wu, R. *Nature* **284**, 426–430 (1980).
5. Capecchi, M.R. *Science* **244**, 1288–1292 (1989).
6. Szostak, J.W. & Wu, R. *Plasmid* **2**, 536–554 (1979).
7. Smolik-Utlant, S. & Petes, T.D. *Molec. cell. Biol.* **3**, 1204–1211 (1983).
8. Orr-Weaver, T.L., Szostak, J.W. & Rethstein, R.J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6354–6358 (1981).
9. Mansour, S.L., Thomas, K.R. & Capecchi, M.R. *Nature* **336**, 348–352 (1988).
10. Robertson, J.S. & Petes, T.D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3550–3554 (1982).

Internal structure of the Earth

Bernard Wood and George Helffrich

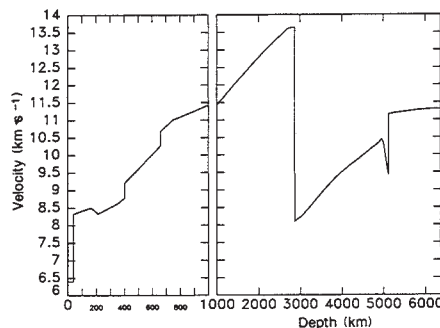
THE chemical and physical properties of the Earth's interior are subjects that influence many areas of the physical sciences but they are, in general, only amenable to indirect study. With the exception of rare inclusions in volcanic rocks there are no samples of the Earth from depths greater than about 50 km, a tiny fraction of its 6,370-km radius. Most information comes from observations of the propagation of seismic waves, with additional constraints being applied by properties such as mass, moment of inertia, electrical conductivity and the presence of a magnetic field. One of the more important traditions in seismology, continued on page 121 of this issue¹ by Peter Shearer, is the characterization of distinct layers in the terrestrial onion. What this layering represents is one of the fundamental questions in Earth sciences.

Seismology as a science dates from the late nineteenth century and it was in 1909 that Mohorovičić discovered² the presence of a velocity discontinuity at about 35 km below the surface in Central Europe. The presence of two (compressional) P-wave and (shear) S-wave arrivals at certain distances from the earthquake indicated that some waves pass down into a high-velocity layer and arrive earlier than those passing through the (relatively slow) upper continental crust. At about the same time³ it was found that P waves passing through the centre of the Earth arrived later than anticipated and that S waves did not arrive at all. By 1914 the presence of a core in the Earth, the outer part of which is liquid (does not transmit S waves), and the gross division of the Earth into crust (0–35 km under continents, 0–10 km under oceans), mantle (35–2,900 km) and core (2,900–6,370 km depth) was well established⁴.

Because it could easily be shown that the overall density of the Earth cannot be obtained by simple self compression (with depth) of the rocks found at the surface, explanations of the three main regions centred on their having different chemical compositions⁵ — a silica-rich crust, a magnesium-iron-silicate mantle and an iron-nickel metal core. These gave about the right density and corresponded to some of the major groups of meteorites thought to represent primordial material of the Solar System. Equation-of-state data for iron (beginning with those of Bridgman⁶) confirm that an iron-rich core containing a few per cent of a light element such as oxygen or sulphur has about the right properties.

Actual experimental results at high pressure lagged behind the early advances in seismology, however, and seismologists

were busy delineating even more Earth structure (see figure). An inner (possibly solid) part to the core was established by Lehmann⁷. As accurate data on the mantle accumulated it became apparent that there is a transition zone of rapidly increasing P- and S-wave velocity between 300 and 900 km depth in the mantle^{8,9} which could not be due to self compression (from 100 to 300 kilobars pressure approximately) and required a change in mineralogical state or chemical composition through this region. Volcanic rocks originating at great depth (recognized because, for example, they bear diamonds) commonly contain fragments of the mantle with about 70 per cent olivine, $(\text{Mg,Fe})_2\text{SiO}_4$, and smaller amounts of pyroxene and garnet.



The variation of compressional P-wave velocity with depth according to refs 9 and 11. The region from 0 to 1,000 km depth has been expanded for clarity.

After Jeffreys's seismological results⁹ were presented, Bernal¹⁰ suggested that, because olivine is probably an important component of the mantle, the rapid increase in velocity between 300 and 900 km depth could be due to the transformation of this mineral to the spinel structure, a transformation that had already been experimentally produced in the analogous germanate. The question of whether or not the discontinuities involve simple isochemical transformations of this type or arise from changes in bulk composition is crucial to understanding the dynamic and chemical history of the Earth. Although the mantle is convecting vigorously, small compositional differences would be sufficient to inhibit convective exchange and isolate the lower regions from most volcanic and other plate tectonic processes.

Refinement of the Jeffreys Earth model has led to replacement of the transition 'zone' by two worldwide seismic discontinuities within the mantle at about 400 and 670 km depth. These each exhibit increases of about six per cent in P-wave velocity with increasing depth and are only a few kilometres wide. Correlations of density with elastic properties indicate

that the 400- and 670-km discontinuities each reflect density increases of about 8 per cent, enough to prevent convective exchange if they are attributable to changes in composition, but not if due to phase transformations¹². Other discontinuities of lesser extent and magnitude have been proposed, most notably those at about 220 km and 520 km that Shearer discusses in this issue¹. By amassing years of long-period (20-second) seismograms from the Global Digital Seismograph Network (GDSN) in a coherent display, he finds that there are distinct phases associated with 400-, 520- and 670-km discontinuities but none associated with a 220-km discontinuity, indicating that the latter is not a globally coherent feature. The magnitude of the discontinuity at 520 km is about half that at 400 km.

Are the discontinuities compositional or due to isochemical phase transformations? The crust-mantle discontinuity (Moho) corresponds to increases in P-wave velocity from about 7 to 8 km s⁻¹ and is generally accepted as being compositional, the mantle being dominated by $(\text{Mg,Fe})_2\text{SiO}_4$ and the crust by more silica-rich minerals. Recent high-pressure experimental studies¹³ have shown that $(\text{Mg,Fe})_2\text{SiO}_4$ olivine does transform to a different structure (modified spinel) at pressures (140 kbar) corresponding to the 400-km discontinuity, the differences in elastic properties and density being of about the right magnitude¹⁴. A second weaker transformation¹⁵ from modified spinel to spinel structure at about 180 kbar may be the explanation for the 520-km discontinuity, and the reaction of spinel structure to $(\text{Mg,Fe})\text{SiO}_3 + (\text{Mg,Fe})\text{O}$ has recently been found to occur at about the right pressure to explain the 670-km discontinuity¹⁵. Thus, the mantle may be of approximately constant composition from 35 to 2,900 km depth and convecting through most of its thickness. □

Bernard Wood and George Helffrich are in the Department of Geology, University of Bristol, Bristol BS8 1RJ, UK.

1. Shearer, P. *Nature* **344**, 121–126 (1990).
2. Mohorovičić, A. *Jahrbuch des meteorologischen Observatoriums in Zagreb (Agram) für das Jahr 1909*, 1–63, (1909).
3. Oldham, R. D. *Quart. J. Geol. Soc. London* **62**, 456–475 (1906).
4. Gutenberg, B. *Göttinger Nachrichten* **1914**, 1–52, (1914).
5. Goldschmidt, V. M. *Z. Elektrochem.* **28**, 411 (1922).
6. Bridgman, P. W. *Proc. Am. Acad. Arts Sci.* **58**, 166–242 (1923).
7. Lehmann, I. *Publ. bureau central seismologique int.* **A14** 87–115 (1936).
8. Byerly, P. *Bull. Seismol. Soc. Am.* **16**, 209–265 (1926).
9. Jeffreys, H. *Mon. Not. R. Astr. Soc., Geophys. Suppl.* **1**, 500–521 (1928).
10. Bernal, J. D. *The Observatory* **59**, 268 (1935).
11. LeFevre, L. V. & Helmberger, D. V. *J. Geophys. Res.* **94**, 17749–17765 (1989).
12. Christensen, U. R. & Yuen, D. A. *J. Geophys. Res.* **89**, 4389–4402 (1984).
13. Katsura, T. & Ito, E. *J. Geophys. Res.* **94**, 15663 (1989).
14. Weidner, D. J. *Geophys. Res. Lett.* **12**, 417 (1985).
15. Ito, E. & Takahashi, E. *J. Geophys. Res.* **94**, 10637–10646 (1989).

COSMOLOGY

Experimental triumph

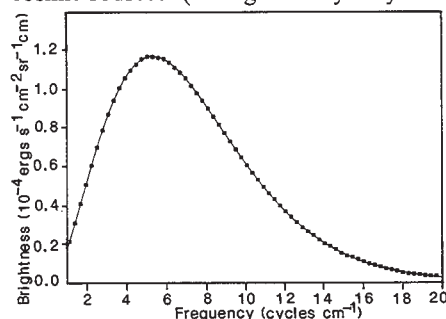
Craig J. Hogan

FOR years, cosmologists have been trying to piece together the history of the Universe — including the history of star formation, galaxy formation and large-scale structure — from fragmentary bits of data. The direct view of the early Universe afforded by viewing its radiation output seems to offer the best hope of probing its history, but except for revealing the existence of the microwave background, experiments have thus far yielded null results. A perfect isotropic blackbody is completely specified by one number, its temperature, and for years experimenters have patiently set increasingly precise upper limits on departures from this ultimate in cosmic ennui. But all this is about to change. An historic moment in cosmology arrived on 13 January this year, with the first announcement of results from the Cosmic Background Explorer (COBE) satellite (see D. Lindley *Nature* 343, 207; 1990). In the first substantial gathering* since the announcement, experimenters and theorists speculated on the exciting developments to expect in cosmology over the next few years.

The by-now-famous measurement of the cosmic blackbody spectrum was made by COBE's FIRAS instrument, the far-infrared absolute spectrophotometer (J. Mather, NASA/Goddard) — “the best-designed instrument I've ever seen” according to D. Wilkinson (Princeton University). The device measures directly the difference of the spectrum of the sky from that of a reference blackbody, over a range of wavelengths from 100 μm to 1 cm. In just nine minutes of data, no deviations were seen, to astonishing accuracy — the maximum allowed fractional deviations from a blackbody, on the basis of these data alone, are less than one per cent. This limit is twenty times smaller than the distortion previously claimed to have been detected during a short rocket flight (see B. Carr *Nature* 334, 650; 1988), which theorists had interpreted as evidence for an era of gigantic pregalactic starbursts. In the course of the next year we can expect the precision of this measurement to improve by one or two orders of magnitude, so that there is hope for a significant detection of a departure from blackbody. Because this spectral region is the repository for just about any ‘waste heat’ that has been deposited by the release of pregalactic energy, we can expect that these beautiful observations will continue to murder many more ugly theories as the data improve.

In fact there will no longer be any place

for significant amounts of cosmic energy to lurk undetected, because COBE's other spectral instrument, DIRBE, the diffuse infrared background experiment, is making maps of the sky at 1° resolution throughout the entire infrared region of the spectrum, all the way from 1 to 300 μm (M. Hauser, NASA/Goddard). The first set of partial maps show many local features already familiar from the Infrared Astronomical Satellite (IRAS), such as a thin bright Galactic plane and a near-infrared zodiacal glow from dust in the Solar System. The idea is that the new spectral information and absolute calibration from DIRBE should allow investigators to model and subtract foreground ‘contaminants’ such as the Galaxy and the zodiacal dust, in order to probe distant cosmic sources. (Along the way they will



Too good to be true? The spectrum of the continuum microwave background radiation as measured by COBE (dots). The full line is the predicted spectrum for a perfect ‘blackbody’ at a temperature of about 2.74 K.

construct a three-dimensional model of the Solar System dust and the most detailed model of Galactic dust so far attempted.) Already the relatively crude rocket data can be used to severely constrain backgrounds in the 100- μm region to a few per cent of the blackbody intensity (A. Lange, University of California, Berkeley). It seems very likely that the signature of primaeval galaxies or even directly redshifted pregalactic starlight might be detected by DIRBE, since it will reach a flux sensitivity of the order of 0.001 of the microwave background intensity, comparable to FIRAS.

The other main instrument on COBE is the differential microwave radiometer (G. Smoot, University of California, Berkeley), which measures the differences in blackbody temperature between points, and therefore maps the anisotropy of the primordial radiation. Here COBE offers a more uniform and complete sky coverage than was available from balloon flights and better sensitivity than the Soviet Relict satellite maps. The instrument has already confirmed the dipole anisotropy arising from the Earth's

motion through the Universe, but investigators are hoping for much more — the detection of intrinsic anisotropies in the Big Bang itself, which would be our first observational data on the inflationary or quantum-gravity era. Two years of integration should smooth out instrumental noise in the maps at the resolution scale (7°) to a root-mean-square fractional temperature fluctuation of just 10^{-5} — a level where many current theories predict a detection.

In spite of its spectacular success, COBE is not the only source of breakthroughs. A race is developing over reconciling the large-scale structure in the galaxy distribution with stringent limits on anisotropy in the background radiation. A variety of programmes using balloons and ground-based telescopes (some of them at the South Pole) are aimed at detecting anisotropy on small angular scales inaccessible to COBE. No new detections were reported, but rigorous statistical limits are being pushed down to record levels, both on scales of arcminutes (C. Lawrence and S. Meyer, California Institute of Technology) and on scales of tens of arcminutes (P. Lubin, University of California, Santa Barbara); fractional temperature variations as high as 2×10^{-5} are now excluded with high confidence. These angular scales are comparable to the angular size of the observed structures in the galaxy distribution, so they ought to reveal anisotropy caused by the very same binding-energy fluctuations that produced the clustering of galaxies and their large-scale velocities — but so far no anisotropy has appeared. The very sophistication of their statistical weapons shows that these teams are circling for the kill.

Taken together, the new anisotropy results are now on the verge (within a factor of two in sensitivity) of offering a serious threat to the most cherished belief of galaxy-formation theory — the idea that large-scale structure arose primarily through gravitational instability (J. R. Bond, University of Toronto; N. Vittorio, University of Rome). Improving data on what the present-day large-scale mass distribution actually is make it harder than ever to sidestep this conclusion (S. D. M. White, University of Arizona). There appears to be no serious experimental obstacle to reaching this interesting level in the next year or two.

In the face of this onslaught of experimental advances, theorists were running for cover in all directions. Several interesting alternatives to gravitational instability of primordial fluctuations were proposed, using hypothetical new physical fields. These fields can causally generate binding-energy fluctuations at late times, after the last interaction of the microwave background with matter. They might do this either by undergoing a straightforward symmetry-breaking transition at

*Aspen Winter Conference on Astrophysics *The Cosmic Background Radiation*, Aspen, Colorado, 21–27 January 1990.

very recent times (W. Press, Harvard University), or by unravelling a unique internal topological knot of a form known as 'texture' (N. Turok, Princeton University; see J. L. Friedman & M. S. Morris *Nature* **343**, 409–410; 1990). It is not clear how these neat ideas will fare when they are worked out as thoroughly as the cosmic-string model (A. Stebbins, University of Toronto; D. Bennett, Princeton University), which is now seriously threatened by both microwave anisotropy constraints and the improving limits on gravitational wave backgrounds.

Astrophysical alternatives to gravitational instability (as opposed to the particle physics alternatives just mentioned) are also in jeopardy. The explosion model is perhaps still just viable, but it generally tends to produce anisotropy which is larger than observed, through the Compton scattering of microwave photons by numerous bubbles of lumpy hot gas along the line of sight (C. Thompson,

California Institute of Technology). The COBE spectral constraints on energy release already imperil this model as well as the mock gravity model, which creates structure by radiation pressure rather than gas pressure. These theories are likely to be either absolutely ruled out or gloriously confirmed when all the COBE results are available.

In the past there has been no real problem in dreaming up new ways of making structure within the improving experimental constraints, but there are real physical limits to what ingenuity can do — and we appear to have reached them. The real excitement for cosmology is that thanks to technical advances we may be coming to the end of the era of null results and will soon have real data on the state of the Universe at early times. □

Craig J. Hogan is at the Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA.

have expected telomere DNA to replicate in a semi-conservative fashion like any other DNA. Years ago, however, it was realized that enzymes that synthesize DNA do so in one direction only, from the 5' to the 3' end of the growing chain. This is fine for the 3' end of a DNA molecule, but it causes problems at the 5' end (see figure). Several inventive solutions to this telomere problem were suggested in the past, but Blackburn's work shows that nature chose a different approach.

The first inkling of something unusual came when telomerase activity was tested on different kinds of DNA molecules. The telomerase of *Tetrahymena* added TTGGGG repeats to the ends of molecules that already terminated in this sequence, but it added the same sequence to molecules with other ends; that is, the TTGGGG sequence was specified by the enzyme, not by the pre-existing telomere. In a similar manner, telomerase extracted from *Oxytricha* or *Euplotes* always added TTTTGGGG sequences^{8,9}. How did the telomerase know which sequences to add? Most enzymes that synthesize DNA or RNA act on a variety of substrates, but the substrate is nearly always the template for new synthesis; that is, sequence specification resides in the substrate, not the enzyme. Blackburn and Greider reasoned that telomerase might contain its own template, and they began looking for a nucleic acid component in their telomerase fractions. They were rewarded by finding a small RNA molecule (159 bases) that co-purified with the *Tetrahymena* enzyme³. When they sequenced this RNA, they knew they were on the right track: at one position within the molecule they found the sequence CAACCCCAA, which is complementary to one-and-a-half repeats added by telomerase. Using a method to cut the RNA in various places while it was still associated with protein, they showed that this sequence was essential for enzymatic activity, and thus likely to be the actual template for new telomeres. Further evidence for the role of telomerase RNA comes from a study, published last month⁹, showing that the RNA from *Euplotes* telomerase contains the sequence CAAAACCCCAA, as predicted from its telomere sequence.

Until now all attempts to reconstitute telomerase activity from separate RNA and protein fractions have failed. Thus Blackburn has been unable to alter the *Tetrahymena* RNA *in vitro* and get *Tetrahymena* telomerase to make 'incorrect' telomeres. As she and her colleagues describe⁴, they circumvented this problem with an elegant genetic experiment. They took advantage of the fact that genes injected into the

TELOMERASE RNA

Tying up the loose ends

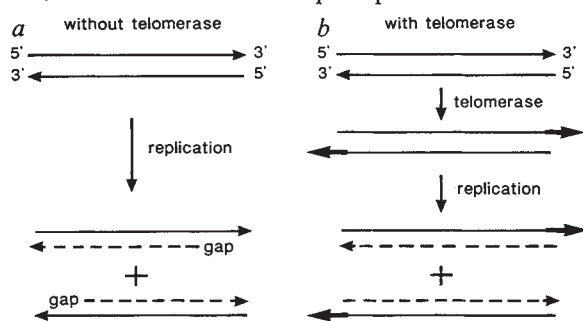
Joseph G. Gall

TELOMERASE is an enzyme that adds simple repeating sequences to the ends of DNA molecules (its name derives from 'telomere', a term long used by cytogeneticists to designate the ends of chromosomes). It was first described from the ciliated protozoan, *Tetrahymena*, by Elizabeth Blackburn and her student Carol Greider¹. Subsequently they found that telomerase is a ribonucleoprotein, whose RNA component is essential for activity^{2,3}. On page 126 of this issue⁴, Blackburn and her colleagues provide conclusive evidence that the RNA encodes the sequences added by the enzyme. This is a considerable advance in the understanding of chromosome structure — and further proof, if more is needed, that novel findings often come from unexpected sources.

If fact, to anyone familiar with the unusual biology of ciliated protozoa, *Tetrahymena* is the logical choice for the study of telomeres. All ciliated protozoa have two nuclei, a micronucleus that contains normal chromosomes, and a macronucleus with fragmented chromosomes. The extent of fragmentation varies from organism to organism, reaching an extreme in *Oxytricha* and *Euplotes*, whose genes are on separate pieces of DNA⁵. Fortunately for molecular biologists, these chromosome fragments are capped with normal telomere sequences, so that DNA isolated from the macronucleus contains many more telomeres than DNA from

any other source. Not surprisingly, telomere sequences were discovered in ciliated protozoa^{6,7}. Many such sequences are now known, and they all consist of simple repeats with one G-rich and one C-rich strand. For instance, the sequence is TTGGGG/AACCCC in *Tetrahymena*, TTTTGGGG/AAAACCCC in *Oxytricha* and *Euplotes*, and TTAGGG/AATCCC in humans and other vertebrates.

The existence of simple repeated DNA



During DNA replication the two old strands (solid lines) are used as templates for new ones (broken lines). Since a strand cannot initiate at the very end of a template, a gap is left at the 5' end of the new strand. *a*, If this gap were not filled in, the molecule would get shorter after each replication. *b*, Telomerase extends the old 3' ends, providing a longer template for the new 5' end.

sequences is not in itself unusual. In most organisms DNA repeats are scattered throughout the chromosomes; they also occur in large blocks near the centromere, the region where the chromosome is attached to the mitotic spindle. What is unusual about telomere repeats is their mode of replication. One might

PLATE TECTONICS

Spinning continents

Dan McKenzie

macronucleus of living protozoa can, under appropriate conditions, partially or totally replace endogenous genes^{10,11}. First they isolated the gene that codes for telomerase RNA (this is not the gene for telomerase protein), and then they mutated it *in vitro* by changing single nucleotides within the putative template sequence. Three different mutated genes were injected into cells. After the cells had grown for a number of generations to allow expression of the injected genes, their telomerase RNA and telomere DNA were examined. In two out of three cases, the injected organisms produced mutated RNA, and their telomeres contained mutated DNA sequences. This experiment demonstrates conclusively that telomerase RNA provides the template for new telomeres.

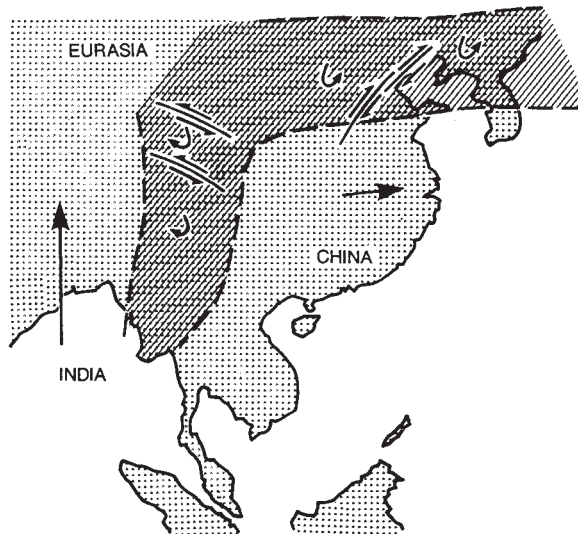
The 1989 Nobel prize in chemistry went to Tom Cech and Sid Altman for their discovery of ribozymes, RNA molecules that function as enzymes. Altman studied RNase P, the RNA component of which is enzymatically active *in vitro*, but which exists as a ribonucleoprotein *in vivo*. Cech studied the remarkable self-splicing pre-ribosomal RNA from *Tetrahymena*; like other ribosomal RNAs, that of *Tetrahymena* is complexed with protein in the organism. Is telomerase a ribozyme? Blackburn calls it a specialized reverse transcriptase, which indeed it is. But if telomerase RNA could be coaxed to function alone *in vitro*, it would become a ribozyme by definition. One can imagine an evolutionary series of RNA molecules beginning with 'pure' ribozymes, proceeding through those that function better with a protein, and ending with ones that now require a protein. Whatever definition is most useful, enzymes with essential RNA components will continue to provide exciting surprises. And given its track record, *Tetrahymena* may be a good place to look for them. □

Joseph G. Gall is in the Embryology Department, The Carnegie Institution, Baltimore, Maryland 21210, USA.

DESPITE the enthusiasm of continental geologists for plate tectonics, it has been clear since the earliest days of the theory that it provides a poor description of continental deformation. The problem is that, unlike oceanic plate boundaries, continents deform in belts up to 2,000 km wide. Although the relative motion of the rigid cores can be accurately described in terms of rotation, all the interesting things that happen in the pliable belts between cannot. This is the fundamental reason why plate tectonics was not discovered much earlier by structural geologists mapping small areas of continents. Only seismologists work with large

south, whereas in Iran the senses are reversed. The effect of all four faults is to move Turkey and Iran away from the Caucasus, where the continental crust is being thickened by the northward movement of Arabia towards Eurasia. Although the first interpretation, by Molnar and Tapponier², of the deformation of central Asia was that the same process was occurring there on a larger scale, this view is now being questioned by England and by Molnar himself³.

As they explain, the reason why they question the earlier view depends on another difference between continental and oceanic tectonics. When deformation



Sketch of the deforming zones between India, China and Eurasia. The arrows show the directions and approximate magnitudes of the velocities with respect to Eurasia.

enough field areas to understand what is going on, and none have a larger one than England and Molnar, who discuss the deformation of half of Eurasia on page 140 of this issue¹.

Once the rules of plate tectonics had been established in the oceans, fault-plane solutions from earthquakes provided a natural way of looking at the present deformation of continents. For some years this enterprise attracted little interest and Peter Molnar had Asia east of 60° E largely to himself, and I worked on the region from the Azores to 60° E.

Throughout the length of the Alpine-Himalayan belt, the plates to the south are moving northwards with respect to Eurasia. It was surprising, therefore, to find that there are very large active strike-slip faults within the deforming zone trending approximately east-west. In the western region these faults occur in pairs, a right-lateral with a left-lateral one. In Turkey the right-lateral one is in the north and the left-lateral one in the

occurs throughout a zone, rather than on a single fault, all the material within the zone can rotate about a vertical axis. A fundamental law of physics is that such rotations are in principle unobservable using the seismic waves radiated from earthquakes. They can, however, be detected by observations of the fossil magnetic fields in rocks. One of the first people to recognize the importance of such rotations was Freund², who mapped an area of eastern Iran that is cut by large north-south right-lateral strike-slip faults to find that some of the motion was taken up by left-lateral movement on many small east-west faults. Beck⁴ and Luyendyk⁵ also found

that palaeomagnetic rotation was common throughout the western margin of the United States, although the fault geometry that they proposed to account for their observations was unconvincing.

Few structural geologists have recognized the importance of this type of deformation. That they should do so is plain: Kissel *et al.*⁶ have shown that extensive regions of central Greece have rotated by 48° in just 5 million years. The recent publication of the proceedings of a conference devoted to this topic⁶ will lead, one hopes, to greater awareness of its importance. Rotations of 90° in 20 million years over belts 200 km wide, are a common feature of continental deformation. Such motions make nonsense of the two-dimensional reconstructions that are still so popular among structural geologists.

England and Molnar propose that the large east-west left-lateral strike-slip faults on the eastern margin of Tibet separate blocks that are rotating clockwise,

- Greider, C.W. & Blackburn, E.H. *Cell* **43**, 405-413 (1985).
- Greider, C.W. & Blackburn, E.H. *Cell* **51**, 887-898 (1987).
- Greider, C.W. & Blackburn, E.H. *Nature* **337**, 331-337 (1989).
- Yu, G.-L., Bradley, J.D., Attardi, L.D. & Blackburn, E.H. *Nature* **344**, 126-132 (1990).
- Prescott, D.M. & Murti, K.G. *Cold Spring Harb. Symp. quant. Biol.* **38**, 609-618 (1973).
- Lawn, R.M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4325-4328 (1977).
- Blackburn, E.H. & Gall, J.G. *J. molec. Biol.* **120**, 33-53 (1978).
- Zahler, A.M. & Prescott, D.M. *Nucleic Acids Res.* **16**, 6953-6972 (1988).
- Shippen-Lentz, D. & Blackburn, E.H. *Science* **247**, 546-552 (1990).
- Tondravi, M.M. & Yao, M.-C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4369-4373 (1986).
- Yu, G.-L., Hasson, M. & Blackburn, E.H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5151-5155 (1988).

and that the whole system forms a huge north-south right-lateral shear zone (see figure). This zone is the largest such structure yet proposed. Their suggestion has several nice features and must surely be basically correct. It allows both the eastward motion of Tibet relative to Asia and India and its northward motion relative to southeastern China to be taken up locally on the eastern margin of the Tibetan Plateau. It also provides a simple explanation for the relative absence of large earthquakes in southeastern China compared with eastern Tibet. No such difference would be expected if the strike-slip faults are taking up the eastern motion of Tibet, but is necessary if they are the consequence of local rotations.

But I think that England and Molnar, like many converts, are too dogmatic in their rejection of the eastern motion of China with respect to Eurasia. The North China Plain is cut by a number of large right-lateral strike-slip faults. The deformation produced by their movement must connect to the western Pacific trenches to the east, and, through the grabens bounding the Ordos Plateau, to the left-lateral strike-slip faults bounding the Tibetan Plateau on its northern margin to the west. As James Jackson has pointed out (personal communication), right-lateral movement on the faults in northeastern China, if they are rotating, can result from left-lateral movement on an east-west trending zone that is even larger than that proposed by England and Molnar.

It is hard to establish whether proposals involving rotating blocks are correct using earthquake data. No such problems affect the direct measurement, with a resolution of a few millimetres, of deformation using Global Positioning System (GPS) satellites. This technique has already been used in central Greece, and has spectacularly confirmed the existence of the rotations proposed to account for the seismic and palaeomagnetic observations there⁷. These early results from GPS suggest that world-wide maps of continental strain rate can be made by direct observation. Such maps would revolutionize our understanding of seismicity and structural geology. □

Dan McKenzie is in the Department of Earth Sciences, Bullard Laboratories, University of Cambridge, Madingley Road, Cambridge CB3 0EZ, UK.

Trailing the itinerant intron

L. A. Grivell

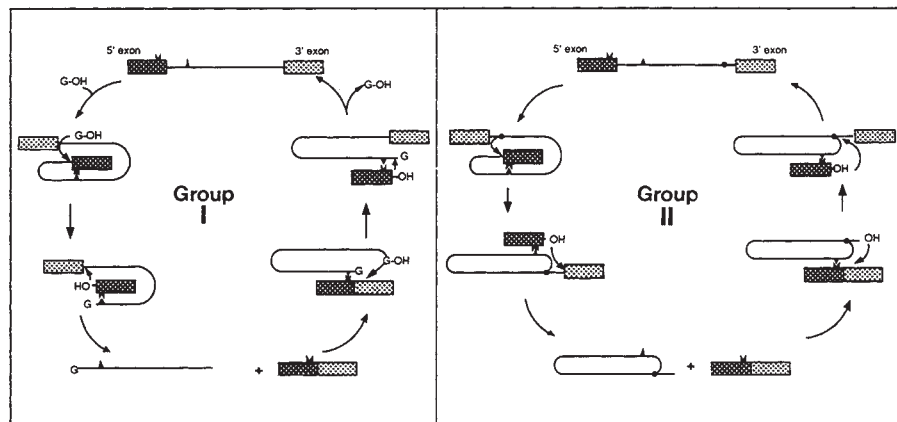
TO ALL but those leading the most sheltered lives, the concept of RNA catalysis has become a familiar one: a growing number of RNA molecules have been found to catalyse a wide variety of reactions involving the breakage and re-joining of phosphodiester bonds in RNA with accuracy and efficiency that were once considered to be the exclusive prerogative of proteins (see ref. 1 for review). Two papers, one in *Nature*², the other in *Cell*³, take the story a step forward.

For introns, the mechanism of self-splicing is pleasingly simple. The intron is excised and the exons are ligated by a series of transesterifications, in which the catalytic molecule allows the formation of one phosphodiester bond by coupling it to the breakage of another. In energetic terms, catalysis occurs because the intron lowers the transition-state energy for the transesterification. Logic dictates that if a catalytic RNA lowers the energetic barrier to intron excision, it must have an equivalent effect on the reverse reaction, which would result in its own insertion

into a molecule containing ligated exons. Persuasive as such logic may seem, however, the low efficiency of the reverse reaction frustrated efforts to obtain direct experimental evidence for reverse splicing of a group I intron until only last year⁴. For group II introns, there have been few firm indications of the feasibility of the reaction.

Now, Augustin *et al.*² and Mörl and Schmelzer³ report that full reversal of a group II splicing reaction is indeed possible. The suspicion that group I and II introns are mobile has gained ground steadily in the past few years. The new results thus open up a route for spreading also of group II introns through a mechanism of reverse splicing, coupled to reverse transcription and recombination.

Self-splicing introns are divided into groups I and II on the basis of a number of criteria, the most important of which are the differences in their predicted secondary structure and the way in which self-splicing occurs *in vitro*^{5,6}. The reactivity of group I introns is dependent on the binding



- Part of the 5' exon which base-pairs with the internal guide (group I) or EBS1 and EBS2 (group II)
- ▲ Internal guide (group I) ▲ EBS1 (group II)
- Intron-internal nucleotide

Self-splicing of group I and group II introns and its reversal. For both groups of intron, forward and reverse splicing reactions occur by a series of transesterifications, after folding of the intron has brought about alignment of 5' and 3' splice junctions. In group I reactions, base-pairing of a part of the 5' exon with an intron sequence known as the internal guide (IGS) is followed by nucleophilic attack on the 5' splice site by a bound guanine nucleotide. The 5' exon thus freed attacks in its turn the 3' splice site, yielding ligated exons and excised intron. In the reverse reaction, the molecule containing ligated exons, or an equivalent site, is attacked by another G residue, this time located at the 3' terminus of the intron. The 3' exon or its equivalent is thus linked to the intron, and the 5' exon or its equivalent is free to attack the 5' end of the intron, liberating the guanine nucleotide bound in the forward reaction. For group II introns, interaction of the 5' exon with the intron occurs by base-pairing with two short sequences (EBS1, EBS2), of which only EBS1 is shown here. The forward reaction is initiated by nucleophilic attack from an intron-internal nucleotide, which subsequently becomes the branchpoint in the resulting intron lariats. The liberated 5' exon attacks in turn the 3' splice site, resulting in exon-exon ligation and intron excision. In the reverse reaction, the intron is aligned with the ligated exons or another potential insertion site, after base-pairing with EBS1. The 3' terminal OH group of the intron is responsible for cleavage, resulting in linkage of the 3' exon, or an equivalent sequence. The 5' exon, or equivalent, subsequently attacks the branchpoint, opening the lariat and allowing integration of the intron.

- England, P.O. & Molnar, P. *Nature* **344**, 140-142 (1990).
- Molnar, P. & Tapponnier, P. *Science* **189**, 419-426 (1975).
- Freund, R. J. *Geol.* **78**, 188-200 (1970).
- Beck, M.E. *Am. J. Sci.* **276**, 694-712 (1976).
- Luyendyk, B.P., Kamerling, M.J. & Terres, R., *Geol. Soc. Am. Bull.* **91**, 211-217 (1980).
- Kissel, C. & Laj, C. *Paleomagnetic Rotations and Continental Deformation* (Kluwer, Dordrecht, 1989).
- Billiris, H. *et al.* *EOS* **70**, 719 (1989).

of a molecule of guanosine or GTP. This molecule subsequently triggers splicing by nucleophilic attack of its 3' OH group on the 5' splice junction (see figure). As might be expected, splicing of these introns is favoured by a high concentration of GTP and concomitant removal of splicing products, whereas the reverse reaction is promoted by absence of GTP and a high concentration of both intron and ligated exons⁴.

By contrast, group II introns are reactive without the intervention of additional nucleotides. Instead, splicing is initiated by attack of the 2' OH of an intron-internal nucleotide on the 5' splice junction. An important consequence is that splicing is accompanied by accumulation of the intron as a lariat, in which the attacking nucleotide closes the loop with a 2'-5' phosphodiester bond (see figure). Splicing of group II introns thus resembles that of nuclear pre-messenger RNAs, prompting speculation that present-day introns in nuclear mRNA-coding genes may be direct descendants of self-splicing entities that first arose in a primitive world dominated by RNA⁷.

What is the evidence for reverse splicing of group II introns? The first tell-tale clue was picked up by Augustin *et al.*³ during their analysis of splicing by a mutant RNA in which the normal 3' splice junction has been replaced by a run of U residues. Compared with a wild-type RNA, the mutant gives rise to novel products, among which is a molecule containing the 5' exon covalently linked to the 5' end of the intron. Such a molecule cannot arise from transesterification during forward splicing, but it does fit into a scheme in which the lariat branchpoint is attacked and opened by reaction with the free 3' OH group of the 5' exon (see figure). Further analysis using labelled RNA substrates shows that full reversal of splicing can be achieved by incubation of ligated exons with mutant intron lariats, and that it results in the formation of small amounts of a molecule that has all the features and reactivity of a normal intron-containing precursor RNA. In wild-type RNA, the forward reaction is apparently favoured by a tendency of the 3' OH of the 5' exon to react with the 3' splice site; in the mutant, the run of U residues may hinder normal nucleophilic attack at this site, thus allowing attack on the lariat branchpoint.

The story does not end there, however. As Mörl and Schmelzer show³, a group II intron can undergo a reverse reaction with a foreign RNA species, provided that the foreign RNA contains a short sequence motif that is complementary to the sequence UCUGUC and that interacts with the 5' exon in a normal splicing reaction (exon-binding site EBS1). The lariat integrates into the foreign RNA immediately downstream of this motif. Although the

sequence of the resulting molecule is as yet unknown, the molecule's ability to carry out self-splicing suggests that correct integration has occurred.

Where do these findings leave us with respect to intron mobility? It seems that group I and II introns can be transposed to new locations (see ref. 8 for review), and indeed the suspicion exists that both have either evolved from or are evolving into mobile genetic elements. For nuclear introns too, their presence in genes at so-called discordant positions⁹ has fuelled speculation on the existence of a continuing process of intron movement. In the case of group I, work on the introns present in T-even phages and yeast mitochondrial DNA has shown that one way in which transposition to cognate sites in intronless genes can be efficiently achieved is by the action of intron-encoded sequence-specific double-strand DNA endonucleases (see ref. 10 for review). These molecules act like highly specific restriction enzymes and cleave the target sequence, allowing intron insertion by double-strand break repair.

Such a mechanism is unlikely to account for the movement of either group of introns to new locations: group II introns lack coding sequences for such enzymes and, for group I introns, the target sequence in the cases so far examined is rather long (18 nucleotides), so that the number of potential new sites is probably limited. By contrast, reverse splicing offers an effective alternative for integration of both types of intron at a new site in RNA, because complementarity to the relatively

short IGS motif (group I) or EBS1 sequence (group II) can be expected to yield many potential sites of integration. The overall efficiency of integration is likely to be low, as the conditions prevailing *in vivo* may further reduce the efficiency of what appears to be a rather inefficient reaction. Action by reverse transcriptase is then required to convert the resulting integrants to complementary DNA which would be able to recombine with genomic DNA. Remarkably, reading frames present in fungal mitochondrial group II introns display similarity to retroviral reverse transcriptases¹¹. The introns themselves may thus provide the necessary activity. This has yet to be demonstrated directly, but the finding that reverse splicing occurs *in vitro* provides one more piece in the experimental jigsaw that one day may allow us to reconstruct the paths that introns take when on the move. □

L. A. Grivell is in the Department of Molecular Cell Biology, University of Amsterdam, Kruislaan 318, 1098SM Amsterdam, The Netherlands.

1. Cech, T.R. *Science* **236**, 1532-1539 (1987).
2. Augustin, S., Müller, M.W. & Schwyen, R.J. *Nature* **343**, 383-386 (1990).
3. Mörl, M. & Schmelzer, C. *Cell* **60**, 629-636 (1990).
4. Woodson, S.A. & Cech, T.R. *Cell* **57**, 335-345 (1989).
5. Cech, T.R. *Gene* **73**, 259-271 (1988).
6. Michel, F., Umesono, K. & Ozeki, H. *Gene* **82**, 5-30 (1989).
7. Sharp, P.A. *Cell* **42**, 397-400 (1985).
8. Lambowitz, A.M. *Cell* **56**, 323-326 (1989).
9. Rogers, J.H. *Trends Genet.* **5**, 213-216 (1989).
10. Dujon, B. *Gene* **82**, 91-114 (1989).
11. Michel, F. & Lang, B.F. *Nature* **316**, 641-643 (1985).

ECOLOGY

Stability from variability

Peter M. Kareiva

IN MOST fields of science, a scattergram that revealed absolutely no relationship between two fundamental variables would mean either that nothing interesting was going on, or that the data were poor. Ecologists often encounter such 'noisy' scattergrams, however, even when the data being plotted concern the relationship between variables one would expect to be linked, such as population density of a host and parasite pressure. Indeed, whereas ecological theory once emphasized the importance of tight relationships between density and mortality rates as a source of the balance in nature, it has become increasingly clear that variation in mortality independent of density is widespread¹⁻³, and under some conditions could be stabilizing^{4,5}. But until the analysis reported by Pacala, Hassell and May⁶ on page 150 of this issue, no one had assessed the frequency with which naturally occurring variation is sufficient, at least theoretically, to stabilize host-

parasitoid systems (or, for that matter, any type of ecological interaction).

The authors analysed data from 21 different host-parasitoid systems and found that, according to a broad class of models, one third of the interactions display spatial variation in the percentage of parasitism that is large enough to ensure stability. The striking finding, especially for those who have traditionally associated a balance of nature with tight negative feedback mechanisms, is that in six of the seven cases identified as potentially stable, the stabilizing mechanism is variation in parasitism that is independent of host density.

The new report describes the latest in a recent wave of theoretical advances concerning the implications of variability for ecological interactions^{7,8}. These advances should be welcomed by field biologists, who have always known that factors such as parasitism vary enormously from place to place, or time to time, but who have not

had the quantitative tools for interpreting such variation. As recently as a decade ago, any field ecologist who recorded widely scattered rates of parasitism bearing no relationship to the density of hosts would probably have shelved the data as useless. We now know that such 'disorder' can be a source of 'order' in species interactions. In particular, if rates of parasitism vary sufficiently from one patch of hosts to another, some patches of hosts can be counted on to escape attack and thereby sustain the host population into the next generation. Before recognition of this point became widespread, explanations for the stability of host-parasitoid interactions were usually sought in inferences about the behaviour of parasitoids, in regulatory mechanisms within the host population itself, or in a rate of parasitism that was tightly coupled to host density⁹.

To merge theoretical ideas about the importance of heterogeneity in parasitism rates with field data, Pacala and colleagues apply what they call the " $CV^2 > 1$ rule". This rule states that, for a broad suite of simple models, if the square of the coefficient of variation in the distribution of searching parasitoids among host patches is larger than one, there is enough variability to stabilize otherwise unstable host-parasitoid interactions. Stability means the presence of an equilibrium density of hosts and parasites that is returned to following perturbation (unless the perturbation is too large); if host-parasitoid interactions lacked this tendency for returning towards some middle-ground of densities, we would probably observe far more frequent extinctions than are usually recorded¹⁰.

To find out how natural interactions fall out with respect to the $CV^2 > 1$ rule, Pacala, Hassell and May took data from 34 studies (21 different interactions), and found that the patch-to-patch variation in parasitism rates is indeed large enough theoretically to stabilize one third of the interactions. An especially interesting facet of the analysis is the apportionment of each interaction's total CV into three components — density-independent; positive density-dependent (parasitism increases with density); and negative density-dependent (parasitism decreases with density). The theoretically stable interactions are found to be so primarily because of density-independent variation; and in the interactions where density dependence is a substantial contributor to CV , it is just as often negative as positive. Thus there is precious little evidence that host populations are tightly regulated by negative feedback in the form of density-dependent parasitism.

The report⁶ is clearly not the final word on heterogeneity in host-parasitoid systems. Although it highlights our need for more data on the spatial patterning of

parasitism rates in natural populations, an exhaustive list of coefficients of variation for one host-parasitoid interaction after another will not be particularly illuminating unless other issues are also addressed. Important questions remain about the robustness of the analysis when applied to alternative models, such as those that allow the searching efficiency of parasites to vary with the density of hosts within a patch. Moreover, the $CV^2 > 1$ rule remains just a hypothesis; it has never been directly tested, or even scrutinized in a correlative manner by relating the temporal variabilities of host-parasitoid interactions to their spatial heterogeneities. We also need a better understanding of the behaviours of parasitoids that generate different spatial patterns of parasitism. Finally, the data gleaned by Pacala *et al.* primarily involve species of economic importance, which surely is not an unbiased sample from which to draw conclusions about stabilizing mechanisms in nature.

Like much good theory, the new analysis causes us to see old data in fresh ways; in particular, rather than viewing environmental heterogeneity and variability as frustrating complications that threaten our ability to make sense of the world, we can now appreciate that this 'noise' in our data may actually confer temporal order on the dynamics of species interactions. It should also stimulate us to put more energy into unravelling the behavioural mechanisms underlying heterogeneity in parasitism and how those behaviours might alter population dynamics. Most importantly, Pacala, Hassell and May have not simply made a contribution to our understanding of host-parasitoid interactions. There is a good chance that the general approach they have adopted will help us understand the role of environmental heterogeneity in a wide range of species interactions, a task that remains one of the central problems in ecology. One has to wonder to what extent other branches of biology gloss over the importance of 'scatter' and variability, in the way ecologists did until the 1980s. □

Peter M. Kareiva is in the Department of Zoology, University of Washington, Seattle, Washington 98195, USA.

1. Strong, D.R. *Trends Ecol. Evol.* **1**, 39–42 (1986).
2. Stiling, P.D. *Ecology* **68**, 844–856 (1987).
3. Walde, S.J. & Murdoch, W.W. *A. Rev. Ecol. Syst.* **33**, 441–466 (1988).
4. Chesson, P.L. & Murdoch, W.W. *Am. Nat.* **127**, 696–715 (1986).
5. Hassell, M.P., May, R.M., Pacala, S.W. & Chesson, P.L. *Am. Nat.* (in the press).
6. Pacala, S.W., Hassell, M.P. & May, R.M. *Nature* **344**, 150–153 (1990).
7. Chesson, P.L. in *Community Ecology* (ed. Hastings, A.) (Springer-Verlag, New York, 1988).
8. Hassell, M.P. & May, R.M. *Ann. Zool. Fenn.* **25**, 5–61 (1988).
9. Hassell, M.P. *The Dynamics of Arthropod Predator–Prey Systems* (Princeton University Press, 1978).
10. Waage, J.K. & Greathead, D.J. *Phil. Trans. R. Soc.* **B318**, 111–128 (1988).

Weight and height

MOST adult animals reach a stable weight no matter how much they eat. Some internal 'weight-stat' adjusts their metabolism to keep it on target. But how does the body know how heavy it is? The obvious answer, says Daedalus, is by direct weighing — the sensing of gravitational pull. He points out that fish seem to lack a weight-stat, and simply keep on growing. Since they are fully buoyed up by the water around them and have zero effective weight, they have no way of judging or controlling their own growth. If land animals could also be buoyed up in this way, their weight-stats should be disabled too.

DREADCO's engineers are now building a special buoyancy chamber to test this theory. It will be filled, not with air, but with some much denser gas. Enlivened with a little oxygen, dense gases like sulphur hexafluoride, halocarbons, even xenon, could support life, even if pressurized. But for total buoyancy, a breathable liquid such as an oxygen-saturated perfluoro-ether would have to be used.

Daedalus expects that an animal buoyed up to half its normal weight will grow to double its normal size. Monster rats and mice might be useful in the laboratory; monster pigs, sheep, chickens and turkeys could revolutionize farming. Poultry farming is already so intensive that a pressurized halocarbon atmosphere would be a minor added complication.

DREADCO's ethologists point out another complication. In a dense enough atmosphere, chickens and turkeys would find themselves able to fly! Indeed, this may be a crucial test of evolutionary theory. A flightless bird like the turkey must be descended fairly recently from flying ancestors. If the software of behaviour is inherited like the hardware of structure, the turkey should still have clumsy and disused flying skills, just like its clumsy and disused wings. Its ancestral habits should be gloriously reawakened in the DREADCO buoyancy chamber.

So Daedalus is asking local zoos to lend him ostriches, penguins, rails, turkeys and other flightless birds, to study their latent flying abilities. Ostriches and penguins, which are used to travelling fast on land or in water, should adapt quite graciously. But turkeys are staid, unadventurous birds. Can the verve and *brío* needed for flight really lie hidden behind that sober and unpromising exterior? The triggering by a small change of atmospheric density of an ancestral but long-suppressed aspect of their personality would be a novel and inspiring experimental triumph. One minor problem is that turkeys are short-sighted. Before they can successfully challenge the heavens, DREADCO's opticians may have to fit them with little spectacles.

David Jones

SIV grows unchanged in human cells

SIR—Adaptation of simian immunodeficiency viruses (SIV) to growth in human cells has been claimed to result in the mutation and selection of virus subtypes¹. Specifically, selection of a premature termination codon in the cytoplasmic domain of the transmembrane protein (TMP) encoded by the *env* gene has been noted and TMP truncation correlated with improved growth in human cells¹⁻⁴.

By contrast, we report here that at least four different isolates of SIV_{agm} from naturally infected African green monkeys (AGM) in our colony consistently express a full-length transmembrane protein with the apparent relative molecular mass of 45,000 (Fig. 1, lanes 2, 3, 5) (ref. 5). These isolates are variants which are easily distinguishable by restriction enzyme mapping and nucleotide sequence divergence. Moreover, a related immunodeficiency

virus (SIV_{abm}) isolated from an African blue monkey caught in the wild also continues to express a full-length gp45 TMP after prolonged growth in human cells (Fig. 1, lane 4).

All our SIV strains have been isolated by cocultivation of phytohaemagglutinin (PHA)-stimulated monkey peripheral blood lymphocytes with human Molt4/8 T-lymphoma cells. For example, after short-term cultivation of 3 weeks, a full-length molecular clone of wild-type SIV_{agm3} (see Fig. 1, lane 5) was isolated and biologically and genetically characterized^{6,7}. The clone was found to be replication-competent *in vitro* and *in vivo*. No premature stop codon in *env* or in any other gene was detected after complete sequence analysis. Western blot analysis of wild-type and molecularly cloned virus after transfection in Molt4/8 cells confirmed the expression of gp45 (Fig. 1, lanes 5 and 6, respectively). These results demonstrate the presence of wild-type viruses with full-length TMPs in their natural host and their successful replication in short-term culture.

Full-length TMPs of our virus isolates were detected even after long-term culture. We were able to show by western blot analysis, the presence of untruncated transmembrane proteins in all SIV_{agm/abm} isolates even after tissue culture for more than a year (Fig. 1).

Two full-length molecular clones of SIV_{agm} have been described^{7,8}. One, termed SIV_{agm} TYO-1, contains a stop codon at a position corresponding to codon 776 in the molecular clone SIV_{agm3} (Fig. 1, lane 7 and Fig. 2)^{5-7,11}. These and all other SIV_{agm} isolates⁹⁻¹¹ are usually grown and replicate well only in Molt4/8 cells. Otherwise, these strains have a very restricted host range, replicating to low titres in CEM and HUT-78 cells and not at all in many other human CD4+ lymphoma lines^{5,10}. Thus, regardless of the unchanged length of their TMPs, replication of SIV_{agm} in human cells is virtually restricted to the Molt4/8 lymphoma line. Infection of seronegative African green monkeys is readily achieved with cell-free preparations of SIV_{agm} with or without

truncated TMPs^{5,6,10}.

All presently available data support the hypothesis that selection against virus substrains predominant *in vivo* occurs only in cell lines that restrict their replication *in vitro*, as observed with SIV_{mac} in several human cell lines. Therefore, perfect replication of SIV_{agm/abm} in Molt4/8 cells or HIV-1 and HIV-2 in human cells lines would impose no drive for selection of adapted mutants. It is easily conceivable that what is observed *in vitro* may also happen *in vivo*. In other words, variants with full-length or truncated TMP may be distributed in an organ-specific manner depending on their potential to replicate in a specific cell type. The studies with SIV_{mac}¹⁻⁴ and SIV_{agm}^{5,8} suggest that genetic information downstream of the premature TMP stop codon may contain a negative regulatory element for HIV/SIV replication in certain host cells.

ALBRECHT WERNER
MICHAEL BAIER
KLAUS CICHUTEK
REINHARD KURTH

Paul-Ehrlich-Institute,
Paul-Ehrlich-Strasse 51-59,
D-6070 Langen 1/Frankfurt, FRG

WALLY BECKER

Department of Medical Virology,
Faculty of Medicine,
University of Stellenbosch,
Tygerberg (7505),
Republic of South Africa

- Hirsch, V.M. *et al.* *Nature* **341**, 573-574 (1989).
- Kodama, T. *et al.* *J. Virol.* **63**, 4709-4714 (1989).
- Chakrabarti, L., Emerman, M., Tiollais, P. & Sonigo, P. *J. Virol.* **63**, 4395-4403 (1989).
- Meyerhans, A. *et al.* *Cell* **58**, 901-910 (1989).
- Kraus, G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 2892-2896 (1989).
- Baier, M. *et al.* *J. Virol.* **63**, 5119-5123 (1989).
- Baier, M., Garber, C., Müller, C., Cichutek, K. & Kurth, R. *Virology* (in the press).
- Fukasawa, M. *et al.* *Nature* **333**, 457-461 (1988).
- Fultz, P.N. *et al.* *J. clin. Microbiol.* **83**, 5286-5290 (1986).
- Ohta, Y. *et al.* *Int. J. Cancer* **41**, 115-122 (1988).
- Kodama, T. *et al.* *J. Virol.* **62**, 4782-4785 (1988).

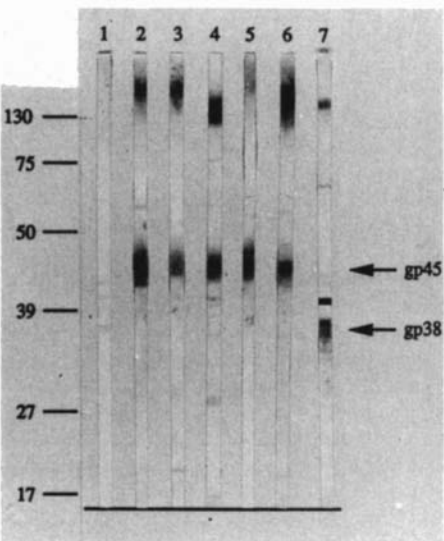


FIG. 1 Western blot analysis of SIV_{agm/abm} antigens. Lanes 1-3: SIV_{agm38} antigen; lane 4: SIV_{abm} antigen; lane 5: SIV_{agm3} wild type antigen; lane 6: molecularly cloned and transfected SIV_{agm3}; lane 7: SIV_{agm} TYO-1 antigen. SIV_{agm} antigen in lanes 2, 6 and 7 were prepared after short-term culture of 4-8 weeks. SIV_{abm} antigen was prepared after 8-months culture. SIV_{agm} antigen in lanes 3 and 5 were prepared after long-term culture of more than a year. Viruses were cultured in Molt4/8 cells. Relative molecular masses (left), in thousands.

CF screening

SIR—Ten Kate (*Nature* **342**, 131; 1989) suggests that carrier screening for cystic fibrosis (CF) should be postponed until 96 per cent of the possible mutations have been detected, on the grounds that so long as a substantial proportion of CF is not associated with the now-identifiable mutation ΔF_{508} , the risk for a couple only one of which is identified as a carrier would still be greater (at 1 in 300) than for unscreened couples (1 in 2,500). We are not so sure.

First, the proportion of CF associated with the ΔF_{508} mutation is likely to be greater in Northern Europe than in North America. Our first impression, after screening 220 chromosomes, is that the proportion of CF caused by this mutation in Britain is more like 80 per cent than the 67 per cent reported by Kerem *et al.* (*Science* **245**, 1073-1080; 1989). If the CF

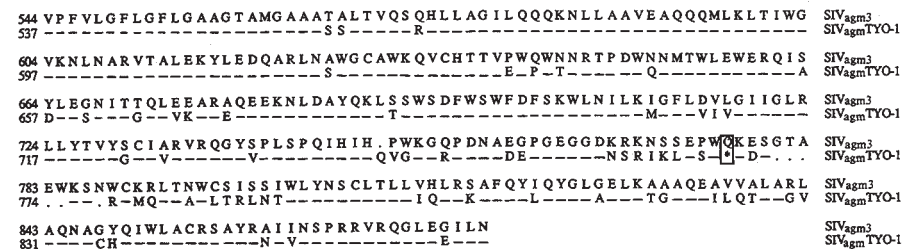


FIG. 2 Sequence comparison between transmembrane proteins of molecularly cloned SIV_{agm3} (refs 5-7) and SIV_{agm} TYO-1 (ref. 11). The dashed line refers to amino-acid conservation with amino-acid numbering corresponding to the positions in the *env*-precursor polypeptides. The premature termination codon of SIV_{agm} TYO-1 is indicated by the asterisk.

carrier rate is 1 in 25 and if 80 per cent of CF genes are ΔF_{508} , the risk to the offspring of a known ΔF_{508} carrier and a partner testing negatively falls to 1 in 500, with at least a proportion destined to have a milder form of the disease.

To compare this risk with that in the unscreened population, as Ten Kate suggests, creates needless anxiety in those who opt for screening. In genetic counselling, we often have to quote risks much greater than for the population as a whole, but which those counselled regard as encouraging. Risks comparable with 1 in 500, for example, obtain in the case of the relatives of those with CF; so far, the great majority have undertaken or continued with pregnancies without any form of prenatal diagnosis.

In the general population, the great majority would test negatively for ΔF_{508} , and the risk factor for couples of such people would fall from 1 in 2,500 to 1 in 62,500. All screening causes extra anxiety for some of those involved, but can it be wise to deny reassurance to the great majority of couples related to known cases of CF?

Were testing to be offered tomorrow to women in early pregnancy, it would perform rather better than the criteria now applied in the avoidance of Down's syndrome, where indicators such as maternal age, low serum alphafetoprotein and high human chorionic gonadotropin

define those with risks greater than 1 in 250 and which, followed by amniocentesis, detect some 60 per cent of Down's syndrome cases.

For Down's syndrome, assuming a population incidence of 1 in 650, in 100,000 pregnancies there would be 5,000 amniocenteses and 92 of 154 cases would be detected (Wald, N., Cuckle, H., and Royston, P., *Lancet* ii, 1362; 1988). With CF, 100,000 pregnancies would require 52 amniocenteses which would detect 26 of 40 cases (arithmetic available on request).

A further benefit of an early start on ΔF_{508} screening is that it would assist with public education, encouraging people to opt for carrier testing before pregnancy and widening their options for the future. Naturally, we acknowledge that, as with other screening for autosomal recessive disease, the need for accurate information on paternity is a consideration of which those to whom testing is offered should be aware. Furthermore we do not underestimate the need that ΔF_{508} screening should be undertaken only when there is a back-up infrastructure of genetic counselling and support.

MAURICE SUPER
MARTIN J. SCHWARZ
I.B. SARDHAWALLA*

*Clinical Genetics Unit,
Biochemical Genetics Unit*,
Royal Manchester Children's Hospital,
Manchester M27 1HA, UK*

promising method in the treatment of patients following a heart attack.

TOSHIFUMI WATANABE
NOBUHIRO SUZUKI*
NORIO SHIMAMOTO
MASAHIKO FUJINO*
AKIRA IMADA

*Biology Research Laboratories and
Tsukuba Laboratories*,
Takeda Chemical Industries,
17-85, Jusohonmachi 2-chome,
Yodogawa-ku, Osaka 532,
Japan*

El Chichòn dust a persistent problem

SIR—Recent discussions on the use of satellite data to derive increases of global sea surface temperatures¹⁻³ have highlighted the importance of understanding all parameters, and their interactions, contributing to the global heat budget. Solid volcanic particles in the stratosphere³ should be included, but the first analyses of satellite data assume that particulates from the El Chichòn volcanic eruptions of 29 March and 3-4 April 1982 interfered only with satellite data collected in the two-year period after the eruptions^{1,2}.

The residence time of aspherical particulates in the stratosphere is probably severely underestimated when calculated on settling rates. A study⁴ of morphology and size distributions of ash particles in the El Chichòn stratospheric cloud, in combination with settling rate calculations for aspherical particles, concluded that the eruptive cloud ceiling reached an altitude of 34-36 km, consistent with lidar and optical polarization observations⁵, or that turbulent mixing of the volcanic cloud significantly enhances the residence time of ash particles in the stratosphere.

In May 1985, a balloon flight from Palestine (Texas) carried a dust collector to sample the stratosphere between 34 and 36 km altitude⁶. Electron microscope analyses of samples (by J. R. Arnold, University of California) show that an important fraction of particulates larger than about 0.5 μm are platy angular shards of silica, plagioclase feldspar and Fe, Mg-silicate grains, and clusters of these grains up to $5 \times 4 \mu\text{m}$ in size. Typically, various amounts of smaller particulates, including sulphuric acid droplets, adhere to these grains.

Settling rate calculations for aspherical particulates show that these large shards and clusters should have settled from 35 km to the tropopause in about one year^{4,7}. The samples also include rare particulates of barite (BaSO_4) and PbO_2 and unidentified submicrometre-sized grains containing iron, nickel, zinc and lead, and occasionally sulphur. Of the plinian volcanic eruptions ejecting par-

Endothelin in myocardial infarction

SIR—There has been a surge of publications on the newly identified vasoconstrictive peptide endothelin since the first report by Yanagisawa *et al.*¹. Although the mechanisms of action and the effects of exogenous endothelin have been widely studied, the pathophysiological roles of endogenous endothelin are still unclear. Recently, we established a monoclonal antibody against endothelin² and an appropriate sandwich enzyme immunoassay system. Using these probes, we have succeeded in demonstrating that endogenous endothelin contributes to the extension of myocardial infarction.

The essence of our findings is as follows. First, treatment with the antibody for endothelin reduces the infarction size in the rat heart. Left-coronary artery ligation (1 h) followed by reperfusion (24 h) induces myocardial infarction in 35 per cent ($n=8$) of the left ventricle in rats. In this infarction model, administration of the antibody significantly reduces the size in a dose dependent manner; 4.5 mg per kg i.v. reduces it to 20 per cent (a 43 per cent reduction in size). Second, plasma and cardiac tissue concentrations of endogenous endothelin are increased to roughly 4-7 times the control value after the coronary ligation-reperfusion procedure: 0.93 to 4.0 pg per ml for the

plasma and 5.0 to 33.1 pg per g tissue for the heart. These results indicate that endogenous endothelin exists in the bloodstream as well as in cardiac tissue under normal conditions, and ischaemic conditions cause an increase in these concentrations, which in turn establishes myocardial infarction.

Several factors have been reported to be involved in establishing myocardial infarction, including free radicals, catecholamines, thromboxanes and various catabolites³. Although there is still controversy whether the elimination of these factors reduces the infarction size, we have demonstrated that neutralizing endogenous endothelin with the antibody results in a smaller size. Therefore, we believe that endothelin is an important member of the family of factors necessary to establish the infarction.

Our findings are also clinically significant. It is beneficial to limit the myocardial infarction, because a larger infarction area may cause severe arrhythmias and haemodynamic depression. Thus, reducing the endogenous endothelin level is a

1. Yanagisawa M. *et al.* *Nature* **332**, 411-415 (1988).
2. Suzuki, N. *et al.* *J. Immun. Meth.* **118**, 245-250 (1989).
3. Reimer, K.A. & Jennings, R.B. *Handbook of Physiology* (eds. Fozzard, H.A. *et al.*) 1133-1201 (Raven, New York, 1986).

ticulates above the tropopause since the mid 1970s, only ash from the El Chichón volcano is (relatively) enriched in these volatile elements⁸. Thus, it is plausible that a significant fraction of particulates collected between 34 and 36 km in altitude during May 1985 derive from those eruptions.

If this is correct, El Chichón volcanic ash particulates persist ($\leq 1 \mu\text{m}$) in the stratosphere at the altitude of the eruptive cloud ceiling for at least three years. These observations, together with abundances of submicrometre volcanic silica particles at a nominal altitude of 18 km⁷, are evidence for longer residence of aspherical particulates and clusters in the stratosphere, if by processes not yet fully appre-

ciated. A bias in satellite data may persist for many years after injection of solid volcanic particles into the stratosphere.

FRANS J. M. RIETMEIJER

Department of Geology,
University of New Mexico,
Albuquerque, New Mexico 87131, USA

1. Strong, A.E. *Nature* **338**, 642–645 (1989).
2. Robock, A.E. *Nature* **341**, 695 (1989).
3. Reynolds, R.W., Folland, C.K. & Parker, D.E. *Nature* **341**, 728–731 (1989).
4. Mackinnon, I.D.R., Gooding, J.L., McKay, D.S. & Clanton, U.S. *J. Volc. Geothermal Res.* **23**, 125–146 (1984).
5. Coulson, K.E., Defoor, T.E. & Deluisi, J. *EOS* **63**, 897 (1982).
6. Testa, J.P., thesis, Univ. of Calif., San Diego, 1988.
7. Rietmeijer, F.J.M. *J. Volc. Geothermal Res.* **34**, 173–184 (1988).
8. Varekamp, J.C., Luhr, J.F. & Prestegard, K.L. *J. Volc. Geothermal Res.* **23**, 39–68 (1984).

Towards the evolution of ribozymes

SIR—Despite current interest in making a functioning ribo-replicase¹ and sequence-specific ribo-endonucleases² (ribozymes), experimentalists have paid less attention to the possibility of intermediate metabolism based on ribozymes^{3–5}. Yet there is evidence, such as the presence of amphiphilic terpenoid lipids with polar RNA fragments in a eubacterium⁶, that organisms based on RNA had a complex metabolism capable of transmethylation, dehydrogenation, carbon-carbon bond-forming reactions and phosphate-ester energy transduction⁷.

Experiments leading to the production of metabolic ribozymes would be of great value, both in science and practically. But I wish to call wider attention to my earlier suggestion^{8,9} that ribozymes with predetermined functions could be produced by *in vitro* replication and selection.

Spiegelman demonstrated that RNAs replicated by the Q β replicase vary considerably and can evolve to predetermined phenotypes by the selection of variants¹⁰. The other source of my ideas is Orgel's suggestion that RNAs with an affinity for small molecules can be selected¹¹. My basic idea is to use catalytic efficiency as the selection criterion. The successful production of catalytic antibodies¹² and the continuing development of Eigen's evolution machine¹³ have helped to put the scheme on a firmer basis.

One could, for example, attempt the following scheme for the *in vitro* evolution of a catalytic ribozyme. RNA variants would be selected on an affinity chromatography column with the stable analogue of the transition-state complex of the reaction to be catalysed, and the selected molecules would then be cloned and subjected to further mutagenesis. This sequence of procedures would be repeated until the molecular population was capable of ribozymal activity.

The use of an analogue of the transition state is borrowed from the techniques for generating catalytic antibodies. A macro-

molecule that can bind to such an analogue is more likely than others to catalyse the corresponding reaction¹². Cloning can be reasonably faithful if, as with protein enzymes, RNAs of a few hundred nucleotides can be catalytic.

The known error-rate of Q β replicase (3×10^{-4} per base replication¹³) allows for accurate cloning. (For comparison, the inaccuracy of the polymerase chain reaction is comparable¹⁴ with 2×10^{-4} misincorporations per nucleotide per cycle). The mutation rate can be controlled by temperature and nucleoside-triphosphate concentration¹³. The whole procedure could be automated using Eigen's evolution reactor, which allows a systematic search of sequence space¹⁵ by first locating weakly catalytic domains and then selecting for perfection within them. That should provide further information about the processes of molecular evolution.

EÖRS SZATHMÁRY

Department of Plant Taxonomy and
Ecology,
Roland Eötvös University,
Kun Béla tér 2,
H-1083 Budapest, Hungary

1. Doudna, J.F. & Szostak, J.W. *Nature* **339**, 519–522 (1989).
2. Haseloff, J. & Gerlach, W.L. *Nature* **334**, 585–591 (1988).
3. White, H.B. III *J. molec. Evol.* **7**, 101–104 (1976).
4. White, H.B. III in *The Pyridine Nucleotide Coenzymes* (eds Everse, J., Anderson, B. & Yu, K.-S.), 1–17 (Academic, New York, 1982).
5. Visser, C.M. *Orig. Life* **14**, 291–300 (1984).
6. Neunlist, S. & Rohmer, M. *Biochem. J.* **228**, 769–771 (1985).
7. Benner, S.A., Ellington, A.D. & Tauer, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7054–7058 (1989).
8. Szathmáry, E. in *Evolution IV Frontiers of Evolution Research* (in Hungarian, ed. Vida, G.) 37–157 (Natura, Budapest, 1984).
9. Szathmáry, E. in *Oxford Surveys in Evol. Biol.* Vol. 6. (eds Harvey, P.H. & Partridge, L.) 169–205 (Oxford Univ. Press, 1989).
10. Spiegelman, S. in *The Neurosciences: Second Study Program* (ed. Schmitt, F.O.) 927–945 (Rockefeller Univ. Press, 1970).
11. Orgel, L.E. *Proc. R. Soc. Lond. B* **205**, 435–442 (1979).
12. Schultz, P.G. *Science* **240**, 426–432 (1988).
13. Eigen, M. *Cold Spring Harb. Symp. quant. Biol.* **52**, 307–319 (1987).
14. Saiki, R.K. et al. *Science* **239**, 487–491 (1988).
15. Maynard Smith, J. *Nature* **225**, 563–564 (1970).

Tacky answer to curious animals

SIR—Research in remote regions often requires that instruments should be left unattended for long periods, when they may be disturbed or damaged by animals¹. High logistical and repair costs are incurred and projects are delayed or may even remain incomplete. Some species have a particular propensity to gnaw at cables. In the course of a programme of ground-temperature measurements in northern Alaska², plastic-sheathed electrical cables were extensively damaged in the neighbourhood of a fox den at Prudhoe Bay. We therefore sought a means of protecting the instrumentation without killing or harming the foxes, which would not have been justifiable.

Many workers in the petroleum fields of North America are natives of the Gulf Coast and eastern Texas, so that hot pepper sauce is abundant in the oil-field dining rooms. Its unpalatability to those not from the Gulf States suggested that it might deter further damage at our field sites.

Replacement cables were suspended, tautened and coated with 'Tabasco' pepper sauce (McIlhenny Co., Louisiana), a layer of clear silicone sealant was applied with a caulking gun, distributed evenly with a sauce-saturated cloth and allowed to dry. While the sealant was still slightly tacky, pepper sauce was applied liberally to its surface.

There was no further damage to our instrumentation in two field seasons. Animals were seen near the instruments, but only shallow toothmarks appeared in the outer layers of sealant. No toothmarks penetrated the cables, which demonstrates the effectiveness and tenacity of the pungency of pepper sauce. A subsequent treatment at Toolik Lake was less effective, perhaps because the treatment was carried out in the rain.

Capsicum frutescens var. *tabasco*, the primary ingredient of pepper sauce we used, appears to be the doyen of peppers, ranking 60,000 to 80,000 Scoville units of hotness. Alternatives such as the Jalapeno and Cayenne peppers rate only 2,500 to 25,000 Scoville units³. The pungent component in hot peppers is the nearly tasteless and odourless substance capsaicin, which stimulates pain receptors in the mouth and throat.

Rozin and Schiller⁴ have discussed the positive affective response to hot peppers

1. Walton, D.W. *Arctic Alpine Res.* **14**, 275–286 (1982).
2. Nelson, F.E., et al. *Arctic* **38**, 310–315 (1985).
3. Andrews, J., *Peppers: the Domesticated Capsicums* (Univ. Texas, Austin, 1984).
4. Rozin, P. & Schiller, D. *Motivation Emotion* **4**, 77–101 (1980).
5. Rozin, P., Gruss, L. & Berk, G. *J. comp. physiol. Psychol.* **93**, 1001–1014 (1979).
6. Rozin, P. & Kennel, K. *Appetite* **4**, 69–77 (1983).
7. Galef, B. *Appetite* **13**, 81–92 (1989).

in humans, but acknowledge that the mechanisms of this 'hedonic shift' are as yet unclear. They find that chili aficionados are not totally desensitized to the irritants in peppers, but rather have grown to welcome the burning sensations they impart. By contrast, Rozin *et al.*⁵ were not able to eliminate the aversion of rats to hot peppers without destroying their capacity to sense chemical irritants, although there have been reports⁶ that domesticated animals respond differently under human

influence, as do rats through the provision of social inducements between individuals⁷.

These considerations do not appear to undermine the potential of pepper products in the protection of cables and other instruments from natural predation.

FREDERICK E. NELSON

Department of Geography,
Rutgers University,
New Brunswick,
New Jersey, 08903, USA

Evolution of tetrapod hearing

SIR—Clack's new data¹ on the stapes of *Acanthostega* are interesting, but there are three objections to the conclusion that a tympanum is absent in primitive tetrapods.

First, theoretical work on the possible function of the Rhipidistian spiracular pouch as a pressure transducing device in water² shows that, if a gas bubble is trapped in the spiracular pouch, it will be in close proximity to the ear so that any pressure change will result in near-field water motion that can act on inner-ear epithelia sensitive to it. This suggests a spiraculum closed by a tympanic membrane.

Second, there is evidence that, in the Coelacanth *Latimeria*, the epithelium of the inner ear is probably homologous to the basilar papilla of tetrapods^{3,4}. Moreover, if the spiracular pouch of *Latimeria*, which is closed by a thin membrane⁵, is filled with gas, the theoretical prediction of van Bergeijk² for Crossopterygians would also apply to *Latimeria*. That would imply that this evolution took place in water and was perfected in parallel with the pressure-transducing mechanisms found in some bony fish^{2,6}.

Both theoretical considerations and the data on *Latimeria* thus favour the view that the spiracle was closed by a tympanic membrane in Crossopterygian fish. The lack of evidence for such a membrane in fossils is of little relevance, given the slim chance that a membrane will fossilize at all.

Finally, there is the question of the transformation of a middle ear able to perceive pressure changes in water into one that performs the same task in air. In principle, this can be done for the aquatic middle ear I have postulated provided that there is a sound-conducting element to transmit the displacement of the tympanic membrane to the inner ear. Van Bergeijk² proposed the hyomandibular bone for this task. The insertion of this ossicle may have happened in some tetrapods without loss of the tympanic membrane. Given that some aspects of the sound-conducting perilymphatic labyrinth and the basilar papilla were already present⁴, it seems reasonable to assume that any change improving impedance matching, even by

as little as one decibel⁷, would minimize the otherwise inevitable loss of 99.9 per cent of all energy impinging on the inner ear⁸ and so help to maintain aquatic hearing in a terrestrial animal.

In summary, I argue that we cannot dismiss the possibility that the otic notch of *Acanthostega* was, as in some Crossopterygian fish, covered by a tympanic membrane and that the slightest benefit in impedance-matching provided even by a massive stapes might have been advantageous⁹. The exciting aspects of *Acanthostega* is the unmistakable connection of the stapes with the otic notch and with the ear as a whole, rather than with the jaws, as is the case with the hyomandibular bone of Crossopterygian fish. Such a changed relationship of the hyomandibular bone is a necessary prerequisite for this bone to start a new function as a sound conducting element in the terrestrial middle ear. In essence, the new finding can be taken to support the old notion of a functional transformation of the hyomandibular bone¹⁰, even if the stapes of *Acanthostega* performed that new function only badly.

BERND FRITZSCH

University of Bielefeld,
Biology Faculty,
Postbox 8640,
4800 Bielefeld 1, FRG

1. Clack, J.A. *Nature* **342**, 425–427 (1989).
2. van Bergeijk, W.A. *Am. Zool.* **6**, 371–377 (1966).
3. Fritsch, B. *Nature* **327**, 153–154 (1987).
4. Fritsch, B. & Wake, M.H. *Zoomorphology* **108**, 210–217 (1988).
5. Fritsch, B. *DFG Reports* **2**, 26–28 (1989).
6. Werner, C.F. *Das Gehörorgan der Wirbeltiere und des Menschen* (Thieme, Stuttgart, 1960).
7. Lewis, E.R. & Lombard, R.E. in *The Evolution of the Amphibian Auditory System* (eds Fritsch, B. *et al.*) 93–123 (Wiley, New York, 1988).
8. Jaslow, A.P., Hetherington, T.E. & Lombard, R.E. in *The Evolution of the Amphibian Auditory System* (eds Fritsch, B. *et al.*) 69–91 (Wiley, New York, 1988).
9. Lombard, R.E. & Bolt, J.R. in *The Evolution of the Amphibian Auditory System* (eds Fritsch, B. *et al.*) 37–67 (Wiley, New York, 1988).
10. Gaupp, E. *Arch. Anat. Anat. Abt., Suppl.* 1–416 (1913).

SIR—Clack (*Nature* **342**, 425–427; 1989) describes a stout element in the otic region of the Devonian tetrapod *Acanthostega* as the earliest-known stapes. The author maintained that there was no tympanum in this form, but rather a spiracular opening, and that the element helped to control palatal and spiracular movements during

ventilation. As the author allows, this function is comparable to that ascribed to the hyomandibular in air-breathing fishes, and in osteolepiform sarcopterygian fishes (the best-supported sister-group of tetrapods). As no auditory function is demonstrated for the element, and as the interpretation presented is that it acted like a fish hyomandibular, we believe that this bone should be called a hyomandibular, not a stapes.

Given the general agreement that the hyomandibular and stapes are homologous, the use of the term hyomandibular in *Acanthostega* more accurately reflects the function of the element and is phylogenetically consistent. In an analogous situation, we do not refer to the quadrate of certain mammal-like reptiles as an incus before it becomes incorporated into the middle ear.

The name stapes is perhaps best restricted to cases in which it is clear that the bone is contributing to the auditory mechanism by supporting a tympanum — the derived condition within tetrapods could then be stated as "presence of a rod-like stapes supporting a tympanum and having an auditory function". The identification of the element in question as a stapes because it occurs in a tetrapod is circular and therefore is not in itself an adequate criterion.

What appears to have been found in *Acanthostega* is not the earliest-known stapes, but retention of the primitive condition — the continued presence of a hyomandibular in a very early and very primitive tetrapod.

MICHAEL D. GOTTFRIED
BRIAN FOREMAN

Museum of Natural History,
University of Kansas,
Lawrence, Kansas 66045–2454, USA

Levitating gyros too good

SIR—The results of Hayasaka and Takeuchi (*Phys. Rev. Lett.* **63**, 2701–2704; 1989) are indeed remarkable, not least because of the minimal scatter in their data (see Fig. 2; S.H. Salter, *Nature* **343**, 509; 1990). If the errors had been accurately estimated, one would expect approximately one-third of the data points to lie more than one standard deviation from the fitted lines. Yet of 31 data points, none is so far away: the probability of observing zero when one expects about ten is $\sim 4 \times 10^{-5}$. Why did neither the authors nor, apparently, the *Physical Review* referees, address this question? I expect most undergraduate physics students to spot such blunders!

A. A. WATSON

Department of Physics,
University of Leeds,
Leeds LS2 9JT, UK

A change of mind

J. Z. Young

Evolution of the Brain: Creation of the Self. By John C. Eccles. Routledge: 1989. Pp.282. £30, \$29.95.

Sir John Eccles says in his epilogue that "My life has been a preparation of almost seventy years for the writing of this book." He can be well satisfied with the result; it is a clear account of his beliefs, which are different from those of most of his fellow physiologists. His theme is the origin of man and his method is to combine the evidence of palaeontology and archaeology with his own immense knowledge of the brain. The result is an extremely compact, well-illustrated and readable account of human evolution. The book's special interest is that it "goes beyond the materialistic concepts of Darwinism", but "only in the last three chapters".

The main part provides a concise and accurate treatment of the fossil evidence for the later stages of human evolution, together with discussion of the changes in the nervous system that must have been needed. Eccles begins with the upright position and bipedal walking, continuing with the origin of language and the relevant changes in the brain. He is not convinced of the capacity for language by apes but believes, with Tobias, that *Homo habilis* was "the initiator of spoken language because of the evidence from endocasts for the existence of the anterior and posterior speech areas".

Eccles discusses the significance of the size of each part of the brain compared with that of monkeys and apes, as recorded by the measurements made by Stephan and his colleagues. Eccles suggests how increase in the size of the pleasure centres of the basolateral nuclei of the amygdala and of the septum could have been associated with the development of monogamous pair-binding, nuclear families and altruism, all providing advantages in food gathering and social life. He dates this change to 3.6 million years ago, on the basis of evidence found by Mary Leakey that a family of three Australopithecines walked together.

Throughout Eccles's accounts of brain function there is little stress on the intrinsic ongoing activity: there is no mention of EEGs or the reticular activating system. One has the curious impression of systems whose input and output are well known but which are not seen in spontaneous action. It is strange that someone who

criticizes "materialistic" and "mechanistic" biologists pays so little attention to the fact that there is constant activity in the brains even of frogs and octopuses, let alone humans, and indeed that activity is an essential feature of life. This lack of a dynamic view may explain why Eccles finds it difficult to understand that the

One may think, on the contrary, that it is very necessary because, on the next page, he claims that mental experiences are essential for vision since "at no stage in the nervous processing can neurons be found that would be instrumental in the eventual reconstruction of the picture, each carrying within itself some particular picture". (Eccles admits that the "rare feature-detection neurons", the face cells, "may be partial exception to this", but one suspects that this was added later.) So he believes that "such an holistic operation apparently cannot be done by the machinery of the cerebral cortex". Other researchers would agree that this problem is difficult but that nevertheless we are

making progress towards a physiological answer. But for Eccles "the emergence of mental experiences in evolution can be understood as providing for integration of the wide diversity of inputs into the brains of highly developed animals".

One may question whether mentality is essential for visual integration, but we all know that we are conscious and can agree with Eccles that "the modern Darwinian theory of evolution is defective in that it does not even recognise the extraordinary problem that is presented". It is a pity that he does not discuss when and how it appeared, as he does later for the divine origin of self-consciousness in individuals.

As a dualist Eccles is anxious to show that mental events produce physical changes, rather than the reverse. For this he describes discharges in the supplementary motor area when a monkey pulls a lever: again, in man there are changes in blood flow in the relevant brain area when there is expectation of a touch on a finger. Such examples are interesting, but

without knowledge of the timing they can equally be taken to show that the brain has produced the mental intention. Indeed the experiments of Libet, not cited here, showed that the brain is active before a person decides on the movement to make a voluntary choice.

Anyhow, no-one will deny that the question of the relation of mentality to the brain is of first importance. The special interest of Eccles' book is that he propounds his precise theory of how the two are related, by the action of "psychons" upon neurons. The essence of this is the hypothesis, ascribed to Margenau, "that mind-brain interaction is *analogous* [Eccles's emphasis] to a probability field of quantum mechanics, which has neither mass nor energy yet can cause effective action at microsites". He believes these sites are at synapses, where "the vesicular

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Thinking big — a comparison of human and chimpanzee brains from *Popular Encyclopaedia* (Cassell, Petter & Galpin, 1860).

brain can make an active search for synthesis: to account for vision of a picture he has to invoke non-material conscious experience.

His treatment of the mind-brain problem is of course the most important part of the book. Unfortunately it is not always clear whether Eccles is writing about consciousness, mentality or human self-awareness. He defines an animal as conscious when it "can exhibit an original behaviour pattern, which can be learned and which also includes a wealth of emotional reactions". This is a wide definition; Eccles then seems to identify it with "mental states", allowing them only to birds and mammals. Unfortunately, he goes on, "for the purpose of this book it is not necessary to study in detail the problem of the origin of consciousness in animal evolution".

Ann Roman

grid provides *the chance (sic)* for the mental intention to select *by choice (sic)* the exocytosis of a vesicle from a bouton". Presumably the mental intention is clever enough to choose a large enough number of psychons to excite (say) a face cell (if he will allow these). Most people will object that it is not legitimate to introduce quantum theory in this way: quantum changes occur, as Eccles himself says, by chance and cannot be directional. However, Eccles here presents his hypothesis in full, available for criticism.

Thus the solution to these problems that he wants us to adopt is that in birds and mammals there is an agency called the "mind" (or "soul" ?) which can perform the holistic feat of "reconstructing the picture". For most biologists it will seem unlikely that some special mental process is needed for vision in birds and mammals but not in reptiles, amphibia or fishes or even in bees and octopuses. Eccles does not provide any indication of when this faculty was conferred.

Indeed, further consideration of the miracle is deferred to the last chapter, "The human person". Here, of course, Eccles is no longer dealing with the mind as involved in physiological processes but with the origin of conscious self-awareness. This he very plausibly identifies with ceremonial burial as performed by Neanderthal man 80,000 years ago. It is indeed a mystery how consciousness arose. Eccles's solution is frankly miraculous. He argues that as all "materialist solutions fail to account for our experienced uniqueness, I am constrained to attribute the uniqueness of the Self or Soul to a supernatural spiritual creation . . . each Soul is a new Divine creation which is implanted into the growing foetus at some time between conception and birth".

Following Eccles's palaeontological studies, one wants to know at what stage of human evolution he believes that this intervention began. But the nearest we get to this is his approval, on the last page, of a quotation from the "Christian Darwinist" Lack, including "in the Christian view, a supernatural event took place at the time of man's first appearance, before which our ancestors were protohuman mammals, and after which, through the divine gift of a soul, they were truly human".

These statements will give comfort to religious fundamentalists all over the world. But the fact that they come from a distinguished student of neurophysiology should not be taken to mean that they are generally, or even widely, accepted. What we can all agree is that these are serious and difficult problems of science and philosophy. We can be grateful to Eccles for his lifelong insistence that we should not neglect them, and for presenting them in such a stimulating book. □

J. Z. Young is at *The Crossroads*, Brill, Buckinghamshire HP18 9TL, UK.



Growing interest — babaco, a highland papaya with a distinctive flavour ("strawberry with a hint of pineapple"), one of the little-known plants of the Andes with potential for worldwide cultivation discussed in *Lost Crops of the Incas* (National Academy Press, \$20).

A day in the life

David Harper

Studying Animal Behaviour: Autobiographies of the Founders. Edited by Donald A. Dewsbury. University of Chicago Press: 1989. Pp. 312. Pbk \$19.95, £15.95.

The Herring Gull's World. By Niko Tinbergen. Lyons and Burford: 1989. Pp. 255. Pbk \$14.95.

MANY scientists believe that the history of their subject is boring. Cynics may argue that this is because having to find out what people actually did and said is all too likely to disrupt scientists' cosy view of unsullied progress. We would far rather ignore the foibles of our heroes and cast some of our intellectual ancestors as villains, than dirty our minds with facts. One of the best cures for a tedious recital of the bare facts of scientific history is an injection of autobiography, preferably well-laced with gossip. The more objective will argue that autobiography is often inaccurate and self-serving, but those, of course, are two of its most interesting features.

Donald Dewsbury persuaded nineteen animal behaviourists to write about themselves. Their accounts first saw the light of day as *Leaders in the Study of Animal Behaviour: Autobiographical Perspectives* (Associated University Press, 1985). The only changes in the revised edition are that the book is now a paperback and that dates of birth have been added to the photographs of the authors.

How on earth does one pick out the nineteen most important scientists in animal behaviour? Dewsbury used a panel of six of the good and the wise to make his

selection. It would be churlish to contest their choice, since any ethologist or comparative psychologist would produce a different list and the panel selected only one of its own number. Many of the more obvious omissions were unfortunately dead (Harlow, Lack, von Frisch), or had already written similar pieces (Beach, Skinner), or declined to contribute (Aronson, Emlen, Thorpe).

The book is valuable because it tells us what nineteen scientists think we should think about their careers. If only Darwin had explained his inordinate fondness for barnacles! An index would have been a pain to produce, but would have greatly enhanced this book. As it is, I now know lots of anecdotes, but cannot remember to whom they should be attributed.

Another book of interest to historians of ethology re-issued as a paperback is *The Herring Gull's World* by Niko Tinbergen, first published in 1960. Like a comforting proportion of the contributions to the Dewsbury volume, this book reveals a deep love of nature and tremendous enthusiasm for science. Tinbergen was responsible for many of the great changes in animal behaviour studies since 1960 and it is surprising how well this book has stood the passage of time. But the discussion of motivation could thoroughly confuse readers new to ethology, and some sections now seem very anecdotal and uncritical. This book was never intended as a comprehensive monograph and readers interested in the bird, rather than the history of ethology, will probably look elsewhere. It remains, however, one of the most modest and inspiring books in ethology. □

David Harper is in the School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, UK.

Tying up the threads

T. L. Killeen

Physics and Chemistry of the Upper Atmosphere. By M. H. Rees. Cambridge University Press: 1989. Pp.289. Hbk £40, \$80; pbk £15, \$27.95.

THE study of the Earth's upper atmosphere, by its very nature, requires an interdisciplinary approach, using tools from many fields of chemistry and physics, including photochemistry, spectroscopy, thermodynamics, fluid mechanics, gas kinetic theory, radiative transfer, plasma physics and optics. Perhaps because of this exciting confluence, available textbooks sometimes over-emphasize one or another area to the detriment of the rest, or provide too much or too little detail for teaching purposes. As a consequence, teachers of university courses have had to select material from various books. M. H. Rees of the University of Alaska has attempted to remedy this deficiency. The task he has set himself is difficult, and one for which there are many possible approaches and pitfalls. In general, his attempt to provide a single comprehensive text is successful, and his book represents an important addition to the literature.

Although the title is all-encompassing, Rees sensibly limits the scope of his book to a discussion of those parts of the upper atmosphere termed the thermosphere and ionosphere. Rees has made many important original contributions in this field and has an obvious and impressive mastery of the theoretical development. Accordingly, his treatment of the relevant physical and chemical processes is thorough and expertly presented, although few fundamental derivations from first principles are given. He discusses key equations and concepts clearly, makes excellent use of recent research, and ends each chapter with a set of thought-provoking (and far-reaching) problems for the reader.

As might be expected from a leading authority on the aurora, the section on the effects on the thermosphere and ionosphere of precipitating particles and solar photons is particularly useful. There is also a valuable chapter on dynamics that is unique in its comprehensive discussion of both the theory and the recent experimental advances. In general, the book works well as a reference text for both senior postgraduate students and specialized researchers, nicely filling a gap in the literature. □

T. L. Killeen is in the Department of Atmospheric, Oceanic and Space Sciences, Space Physics Research Laboratory, University of Michigan, Ann Arbor, Michigan 48109, USA.

Happy birthday

Joël Janin

Molecular Biology: A Selection of Papers. Compiled by S. Brenner. Academic: 1989. Pp. 622. Pbk £15.95, \$29.95.

WAS Balzac the first to make the age of thirty a subject for literary celebration? His *Femme de Trente Ans* describes a lady who has been through the pains and pleasures of youth embarking on a new stage of her life as she turns thirty. The *Journal of Molecular Biology*, known as *JMB* to its many friends, was founded by John Kendrew in 1959 and has now reached balzacian age. Its anniversary is celebrated in a book where some of the best papers are collected. Compiled by the present editor-in-chief, Sydney Brenner, it extends from 1959 to 1975 and covers 100 volumes of the journal. The selection of 38 papers among several thousands must have been very difficult, for which Brenner takes full responsibility. As the texts included are superb and well-worth reading even a quarter of century after they appeared, he can be blamed only for the inevitable omissions.

What was molecular biology in the 1960s and early 1970s? A selection of *JMB* papers should answer this question. Undoubtedly, molecular biology was then the biology of *Escherichia coli* and its phages, and more than half of the papers reprinted here deal with their genetics. The big move towards eukaryotes had just

New in paperback

■ *Neural Darwinism: The Theory of Neuronal Group Selection* is published in the United Kingdom today (8 March) in paperback by Oxford University Press, price £9.95. For review see *Nature* 331, 571; 1988. Edelman's most recent book, *The Remembered Present*, was reviewed by Edward S. Reed in *Nature* 343, 603; 15 February 1990.

■ Two books on quanta have recently been issued in paperback. *Are Quanta Real*, by J.M. Jauch, contains a new foreword by Douglas R. Hofstadter. First published in 1973, Jauch's book is a "Galilean dialogue", using three characters to represent different viewpoints of the physics and philosophy of quantum mechanics. Publishers are Indiana University Press, price is \$7.95. *Atoms and Quanta* by Daphne F. Jackson is an undergraduate introduction to quantum mechanics. Published by Surrey University Press, price £15.95; \$29.95.

■ Two new publications in the IRL Press series *Frontiers in Molecular Biology*, edited by D.M. Glover and B.D. Haines, are *Oncogenes*, published on 18 January, and *Molecular Neurobiology*, published on 15 February. Prices are £18; \$38 for each volume.

■ In *The Threatening Desert*, Alan Grainger discusses causes of desertification and strategies for its control. The publisher is Earthscan, price is £9.95.

started in 1975, where the selection stops. More surprisingly, the favourite molecule of molecular biology seems to have been transfer DNA rather than a protein or even DNA. Full papers are devoted to two tRNAs, and when the DNA era started, the first paper is on the synthesis of the alanine tRNA gene by Khorana and collaborators in 1972. No protein is given much credit, save perhaps alpha-chymotrypsin, whose three-dimensional structure and mechanism were described by Blow and collaborators in 1968.

The absence of immunology and of developmental biology reminds us that these fields turned molecular only in the late 1970s. Structural studies, on the other hand, have been fundamental to molecular biology from the start. The book includes several major papers in electron microscopy, covering progress of the technique from early work on muscle and viruses to the advent of high-resolution models of unstained specimens in 1975. Good reprints of electron micrographs, which still strike the eye, are one of the best features of the book. By contrast, the role *JMB* played in protein crystallography is hardly apparent. The journal has been (and still is) the chief journal of protein crystallography. Crystallographers were actively publishing protein structures in the seventies, yet only one paper is reprinted here, that on alpha-chymotrypsin. Is this because crystallography papers carried none of the aesthetic and little of the scientific value of the structures they described? They certainly used to be illustrated by rather unintelligible diagrams in black and white, with stereo pairs that were difficult to see in three dimensions. Protein structures became popular only with the advent of sophisticated computer graphics.

The book makes it clear that *JMB* in the 1960s was very much a family affair. The journal's home was in Cambridge at the MRC Laboratory of Molecular Biology, where John Kendrew and Sydney Brenner worked together with many of its most prestigious authors. No less than 16 of the 38 papers selected by Brenner originate from the Cambridge lab. They include Nobel prizewinning reports by Aaron Klug on the electron microscopy of viruses and by Fred Sanger on nucleic-acid sequencing, and also four fine papers by Brenner himself and his collaborators. Some of us may remember that molecular biology had a few strongholds outside Cambridge. Only two are prominent in the selection: the Pasteur Institute, with several papers by Jacob and Monod, and the Biological Laboratories at Harvard, with Meselson, Ptashne and others. The rest of the world is too sparsely represented. Still, I was glad to find Cairns' famous images of the *E. coli* chromosome among the highlights of 1963.

Brenner's highly personal selection

leaves an impression of the early years of molecular biology which is probably truthful for Europe, at least. Molecular biology still had to claim the right of independence from biochemistry and genetics. Although major biochemical journals were published in the United States, *JMB* was a European adventure. As molecular biology became an established field, its scope grew to absorb cell biology and development. New journals were created and *JMB* had to face strong competition after a long period of near-monopoly. Nowadays, authors are still attracted by its reputation, by the high technical quality of

the typography and by the convenience of having almost no limitation in length. Yet high subscription costs and slow publication are serious drawbacks. *JMB* has lost some of its audience even as its scientific excellence is maintained. Is that a sign that old age starts at thirty for a scientific journal, or shall we trust Balzac's word that "a lady of thirty has irresistible appeal on a young man", and count on the new generation of molecular biologists to match the one that created *JMB*? □

Joël Janin is Professor of Biophysics at the Université Paris-Sud, 91405 Orsay, France.

Ancient bees and bronzes

Clive Gamble

Economic Prehistory. By Grahame Clark. Cambridge University Press: 1989. Pp. 638. £50, \$90.

Economic Anthropology. Edited by Stuart Plattner. Stanford University Press: 1989. Pp. 487. Hbk \$49.50, £29.36; pbk \$16.95, £10.05.

NO-ONE would contest Grahame Clark's pre-eminence as the founder of economic and social approaches to the study of our prehistoric past. But neither would anyone be surprised that his influence has not spilled over into even the disciplines most close to archaeology, as shown by the absence of citation to his works in Stuart Plattner's edited volume on *Economic Anthropology*.

For the archaeologist, the 32 papers in *Economic Prehistory* — spanning more than 40 years — will affirm again that Clark single-handedly redirected the subject away from sterile museology, where the past was treated like stamp collecting, and towards the study of what people did. Clark put people back into the past by concentrating on hitherto neglected topics in the prehistory of northwest Europe, such as fishing, fowling and whale hunting, as well as using the pollen evidence for forest clearance and linking this to farming settlement. In the earliest paper reprinted here, "Bees in Antiquity" (1942), Clark summed up his view in a lapidary phrase "which tell us more, the bees or the bronzes?", one he returned to in his retrospective (1974). This 'question-begging question' encapsulates Clark's ecological approach, with its emphasis on the inter-relation between environment and society that links honey with land use and beeswax with bronze-casting. Such a holistic approach was summarized in diagrams that long predated the systems analysis of prehistoric societies. Clark's ability to connect evidence from different spheres of activity

and present them through his classic case studies of hunting camps, such as his own excavations at Star Carr, or sealing villages around the Baltic, established the alternative to museology as well as the sterility of typological analysis for its own sake.

These studies mark Clark out as an optimist in terms of what can be done with prehistoric data. Significantly, his book *World Prehistory* (1961) was dedicated to two other archaeological optimists, Childe and Crawford, who with Clark lay claim to being by far the most important British prehistorians of the century because of their visions of what the subject could contribute to understanding human history. Clark built his on economic foundations which, he tells us, were never the key but merely a useful approach in attaining a wider goal. What this goal might be emerges in the 13 papers grouped here under the headings of "World Prehistory" and "Archaeology and Society", mostly written between 1960 and 1980.

Clark claims that "the substance of prehistory is no less than the emergence and self-realisation of humanity itself" (1968) and recognized archaeology's appeal in telling us about our identity (1980). Such statements do not fit easily with archaeology's present mood of introspection and particularism. But as Robin Dennell has pointed out, Clark is one of the few prehistorians who have tried to tackle the question of archaeology's relevance. Although we may not accept his championing of cultural diversity and the retention of separate cultures as a bulwark against what he sees as dangerous contemporary currents of a homogenizing culture, or the value of social élites as patrons of great art, his willingness to tackle contentious issues has to be acknowledged. At the same time, in his penetrating essay on human diversity (1979), Clark shows he is well aware that the esteem for archaeology flows from "its ability to dignify and validate the status quo by investing it with the sanction of antiquity" and that as a result archaeologists, unlike anthropologists, conform rather than protest. Although accurately describing the social role of most archaeologists, Clark's work has not

provided the springboard for any radical critique although protest, as shown by recent battles over reburial issues, the Rose theatre in London, and the corruption of heritage, academic freedom and apartheid, is now increasingly part of the archaeological agenda.

Clark's papers reveal two main events which shaped his thinking: the wartime mechanization of the countryside and the loss of folk traditions (1951); and the loss of the British Empire, indicated by his papers on Australia (1968) and world prehistory. The first was a legacy from O.G.S. Crawford: knowledge of how traditional agriculture worked now had to be learnt from books and passed on to students. What was once common knowledge now became articulate explanation, so elegantly contained in *Prehistoric Europe: the Economic Basis* (1952). The second resulted in Clark's global empire of Cambridge graduates equipped with his vision and techniques which were enhanced wherever they also discovered that, outside the joint Cambridge faculty, archaeology could be combined with anthropology in more than just name.

What many of these students know, but what is missing in this remarkable set of papers, and even in Clark's invited lectures to learned US societies, is that the intellectual base of the discipline has shifted across the Atlantic. The present and indeed postwar dominance of anthropology by researchers in the United States is revealed in the excellent collection of thematic essays in Stuart Plattner's volume. Substantivist and formalist, marxist and development economics are thoroughly described and discussed in what will become a standard and much-needed undergraduate text. Wider issues of trade, markets, the role of women, the informal economy and common property, as well as chapters on hunters and gatherers and horticulturalists, are covered in 15 commissioned papers. The often seamless transition from economic anthropology to development economics underlies the embeddedness of the discipline in the present, retaining its identity only through its insistence on analytical scale within an otherwise all-comers paradigm.

The aim of the authors of these papers is to set out what economic anthropology knows about the world, and this is achieved with great success and various perspectives. As a result it is not difficult to see why Clark's present economic values and social positions, justified by his reading of the past, are ignored in these essays. For, as Frank Cancian concludes in his contribution on peasant communities, "there are many reasons to think that active peasants and their descendants will put their own stamp on the future". □

Clive Gamble is in the Department of Archaeology, University of Southampton, Southampton SO9 5NH, UK.

Seismic imaging of upper-mantle structure with new evidence for a 520-km discontinuity

Peter M. Shearer

Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA

Stacked images of long-period seismic data reveal many phases, some not previously observed, which are caused by reflections and conversions at discontinuities in the Earth's upper mantle. These images support the existence of discontinuities at 410, 520 and 660 km depth, but no evidence is found for a significant interface at 220 km.

SEISMIC velocities in the Earth's upper mantle increase more rapidly with depth than is to be expected from the increasing pressure alone, indicating that phase changes or differences in rock chemistry must be occurring. At some depths seismic properties change very rapidly, and relatively sharp discontinuities must be present. Although many possible discontinuity depths have been proposed, most recent models of upper-mantle structure contain discontinuities only near 400 and 670 km depth. An understanding of these seismic observations has been sought from high-pressure experiments in mineral physics, which suggest that the 400-km discontinuity is caused by the phase change between α - and β -olivine¹, and the 670-km discontinuity can be explained by the transformation γ -olivine (spinel) $\rightarrow$ perovskite + magnesio-wüstite², although some researchers argue for chemical changes at these discontinuities³. These experiments have direct implications for the continuing controversy over whether the entire mantle has a relatively uniform composition in which discontinuities represent phase changes alone, or whether it is stratified into layers of different chemical composition. Seismology may help to resolve this dispute by providing more precise information on the strength, sharpness and possible topography of these discontinuities.

In addition to the well established major discontinuities near 400 and 670 km, there is some seismic evidence for other upper-mantle discontinuities. Numerous studies⁴⁻⁸ have found evidence for a discontinuity or steep velocity increase near 220 km, leading some to suggest that a 220-km discontinuity is an important global feature which occurs beneath both oceans and continents⁷. Proposed explanations for a 220-km discontinuity have included a composition change between garnet lherzolite and eclogite⁷ and the existence of an anisotropic layer caused by preferred alignment of olivine crystals⁸. But many recent seismic studies^{9,10} of upper-mantle structure are inconsistent with a significant discontinuity at 220 km, so it is unclear how globally coherent such a feature can be. A small number of studies have also suggested a discontinuity or enhanced velocity gradient at 520 km depth^{5,11-17} which may be associated with the β -olivine $\rightarrow$ spinel phase change¹. Discontinuities at various other depths have been proposed as well^{5,11,12}, but with little consensus between different studies regarding their occurrence.

Seismic evidence for discontinuities in the upper mantle began with the recognition of a sharp bend in the P wave travel-time curve near 20° range^{18,19}, and has continued with detailed travel-time and synthetic seismogram analyses of both P and S wave phases^{9,10,20}. Other evidence for upper-mantle discontinuities

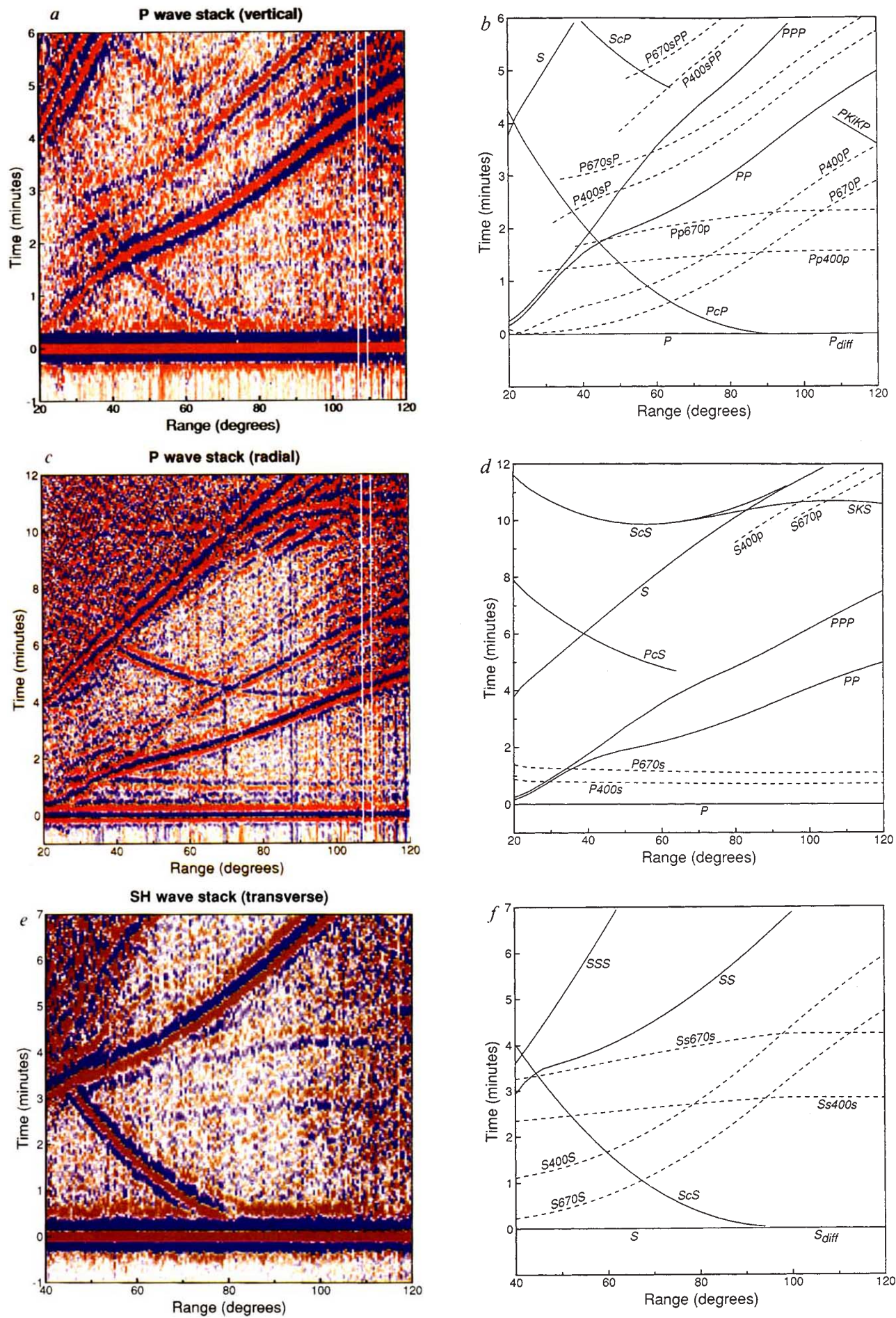
has come from observations of secondary arrivals resulting from topside reflections^{11,21-26}, bottomside reflections^{21,26-32} and phase conversions between P and S at the discontinuities³³⁻⁴¹. Again, many different discontinuity depths have been suggested by these observations, with general agreement occurring only for the 400- and 670-km discontinuities. Most of these studies of secondary arrivals have concentrated on examining data from a relatively small number of earthquakes and seismic stations, often with poor signal-to-noise characteristics, and it is not always clear which features of the models are global in nature and which may be peculiar to a particular region or data set.

The current availability of large seismic data sets in digital form makes it practical to 'stack' and average thousands of seismograms to remove the effects of local structure and enhance the visibility of weak global seismic phases. Here I present the results of stacking five years of long-period seismic data from the Global Digital Seismograph Network (GDSN). Images resulting from these stacks clearly show many phases related to upper-mantle discontinuities; some of these phases have not previously been observed. Strong phases are seen associated with discontinuities at apparent depths of about 410 and 660 km. Depths calculated for individual stations agree to within 20 km of these values, suggesting a possible limit to the amplitude of any topography on these discontinuities. In addition, these images appear to show two seismic phases from a discontinuity at 520 km depth, thus supporting the results of earlier studies which have indicated structure at or near this depth^{5,11-17}. No evidence is found for a 220-km discontinuity, although phases associated with such a discontinuity should be visible in some of the stacked images if the 220-km discontinuity were a significant global feature.

Stacking long-period seismograms

For more than ten years, seismograms from the Global Digital Seismograph Network (GDSN) have been available through the National Earthquake Information Center at the United States Geological Survey. This study uses data only from the five years (1980-1984) that are currently available on optical disk (CD-ROM), the most convenient form for automatic processing of large parts of the data set. These five years contain records for 1,827 earthquakes and nuclear explosions, with an average of about 21 stations per event, for a total of 38,578 event-station pairs.

My analysis is limited to the long-period GDSN data, which record three components (vertical and two horizontals) sampled at 1 Hz, filtered such that the seismic pulses recorded have a dominant period of about 20 s. Long-period seismic data are particularly suited for stacking (the addition of different seismic records to yield a single composite record with higher signal-to-noise ratio) because of the coherence and simplicity of typical long-period waveforms. The stacking procedure works in the following way: (1) The analysis is restricted to events shallower than 50 km to eliminate complications owing to depth phases (surface reflections near the source). (2) The amplitude and start time of each seismogram are normalized with respect to a reference seismic phase. This procedure increases the coherence



◀ FIG. 1 Stacked colour images of long-period GDSN data compared with travel-times curves for primary seismic phases (solid) and upper-mantle discontinuity phases (dashed). Positive amplitudes are shown in red, negative amplitudes in blue, with the scale ranging up to 0.1 of the maximum amplitude of the reference phase. The stacks include: *a*, *b*, P wave reference phase, vertical-component stack; *c*, *d*, P wave reference phase, radial-component stack; *e*, *f*, SH wave reference phase, transverse-component stack. Note the apparent phases visible in *a* between Pp400p and Pp670p, and in *e* between Ss400s and Ss670s, representing possible reflections off a 520-km discontinuity.

of the stack and reduces the effects of differing source functions and instrument responses and errors in earthquake depths and origin times. Both P wave and SH wave reference arrivals are used in the stacks shown here. (3) The maximum amplitude in the seismogram is found within 60-s windows both preceding and following the theoretical arrival time of the reference phase. These amplitudes are used as a measure of the size of the 'noise' level before the reference phase arrival and of the 'signal' contained in the reference phase. The seismogram is discarded if the signal-to-noise ratio is less than 10. This ensures that only the higher-quality data are used in the stack. (4) The amplitude scaling and start time of each seismogram are adjusted such that the maximum amplitude of the reference phase is unity at zero time. The polarity of the seismogram is switched if necessary. This has the effect of aligning the peaks in the reference phase for all the seismograms in the stack. In general, the waveform shapes in the reference pulses will not be the same, mainly because of differences in event depths. When applied to dozens of events, however, this stacking procedure results in an effective source function that is peaked as sharply as possible. (5) Values obtained from each trace are averaged within bins of range and travel time to produce a stacked record section combining all the data.

These stacked images are seen most clearly on displays in which positive and negative amplitudes are represented with different colours so that the polarity and phase of the various seismic arrivals can be seen directly.

P wave stack (vertical component)

Figure 1*a* shows the result of aligning the P wave arrival and stacking on the vertical component at ranges between 20° and 120° and travel times of up to 6 min following the P wave arrival. The data are binned and averaged in increments of 0.5° in range and 3 s in time. 2,332 seismograms are stacked from 35 stations, for an average of about 12 seismograms per range bin. Positive amplitudes are shown in red, negative amplitudes in blue. Note that the P wave arrival has been aligned at zero time, and the maximum amplitudes shown are only one tenth the size of the P wave amplitude. This enhances the appearance of the weaker seismic phases, at the expense of 'clipping' the strong arrivals. At ranges beyond about 95°, the P wave has been diffracted around the core-mantle boundary, corresponding to the ray path labelled P_{diff} in Fig. 2. P_{diff} is a weaker phase than direct P, reducing the number of good-quality records available for stacking. This causes the stack to become noticeably noisier at ranges beyond 100°, because of the smaller numbers of traces at these ranges. The two white vertical lines towards the right of the plot are the result of the absence of data within two individual range bins.

Figure 1*b* shows theoretical differential travel times calculated by tracing rays through PREM⁴², a reference Earth model. Major phases are shown as solid lines; these include P, PP, PPP, PcP, PKiKP, S and ScP (Fig. 2 shows the geometry for many of these phases). The phase of a particular seismic arrival can be identified by its colour pattern. For example, the blue-red-blue pattern of the core-reflected phase PcP indicates that it has not been phase-shifted relative to the direct P arrival. By contrast, the red-blue-red-blue pattern of the surface-reflected phase PP shows the $\pi/2$ phase shift (the Hilbert transform) known to occur for this phase⁴³.

The dashed lines show some of the upper-mantle discontinuity phases discussed in this paper. Figure 3 shows examples of ray paths for these phases. Perhaps most visible are the two phases P400sP and P670sP, which follow PP and contain a single S leg between the 400- and 670-km discontinuities and the surface ('400' and '670' will be used in the phase names and discussion, although the actual discontinuity depths may be slightly different). These phases can be followed from a range of 30° to 120°. Figure 3*a* shows only one of the four possible geometries for these phases; the S leg could replace any of the P legs shown and the same travel time would result. Analogous to these phases are the arrivals that follow PPP (namely, P400sPP and P670sPP) which also contain a single S leg between the discontinuities and the surface. These phases are visible beyond about 50°.

Weak phases can be seen following the direct P wave arrival by about 1.5 to 2.5 min at ranges between 70° and 120°. These are P wave multiples (Pp400p and Pp670p) which result from topside reflections off upper-mantle discontinuities²¹⁻²⁵. Notice that they are phase-reversed relative to the P wave, a result of the surface reflection. An apparent phase can be seen between the Pp400p and Pp670p arrivals. The position of this phase corresponds to a topside reflection from a 520-km discontinuity (that is, to Pp520p). An analogous phase appears in the SH wave stack and will be discussed later. Also shown in Fig. 1*b* are travel-time curves for the bottomside reflected phases P400P and P670P which precede PP (refs 21, 25, 28, 29). These phases are visible only marginally in the stacked image, most clearly between 90° and 110°. Noticeably absent in Fig. 1*a* is any indication of a multiple from a 220-km discontinuity (Pp220p), which would appear between P and Pp400p.

P wave stack (radial component)

Figure 1*c* shows the result of stacking on the radial horizontal seismometer component after aligning the P wave arrival as in the previous example. The binning increments are the same but results are displayed out to 12 min to show some of the details near the S wave arrival. The maximum amplitude shown is 0.1 of the P wave amplitude. Figure 1*d* shows differential travel times for the phases visible in this stack. These include P, PP, PPP, PcS, S, ScS and SKS. Note that PcS is visible out to 90°, well beyond the ray theoretical limit for this phase. At ranges between 65° and 90°, this phase is actually PcS_{diff} , with the P wave diffracting around the core-mantle boundary (see Fig. 2).

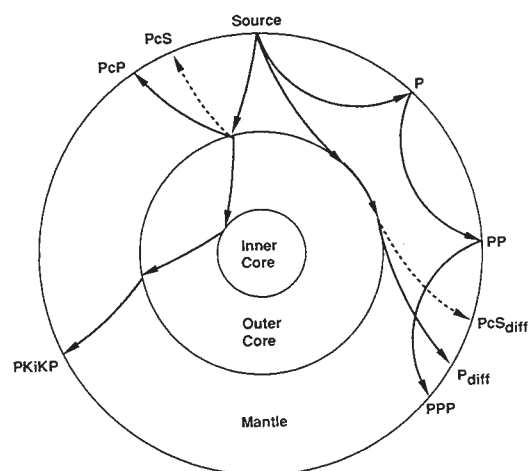


FIG. 2 Schematic of the major global seismic phases seen in Fig. 1. The phases are associated with interactions with three discontinuities: the surface, the core-mantle boundary, and the inner-core boundary. Compressional waves (P) are shown as solid lines; shear waves (S) are shown as dashed lines. Reflections are indicated with lower-case letters; for example, PcP is the P wave which is reflected off the core-mantle boundary, and PKiKP is reflected from the inner-core boundary. P_{diff} and PcS_{diff} are phases that have been diffracted around the core-mantle boundary.

Preceding the S arrival are two phases, S400p and S670p, that result from SV to P conversions at the 400-km and 670-km discontinuities beneath the station (see Fig. 3)^{34,37,38,41}. These arrivals may also contain energy from P to SV conversions at discontinuities below the events (p400S and p670S); these phases arrive at the same time as the SV-to-P converted phases, and both sets of arrivals are seen most clearly on the radial component³⁷. S400p and S670p can be seen at ranges as small as 60° and 80° respectively, much closer than the ray theoretical limits shown in Fig. 1d but in agreement with the results of synthetic seismogram modelling³⁴.

Two phases—P400s and P670s—can be seen clearly as blue streaks in Fig. 1c, which follow the direct P by about 40 s to 90 s. These result from P to SV conversions at discontinuities below the recording stations^{33,39,40}. The time separation between these phases is too short to permit a 520-km discontinuity to be seen, because the dominant period of the pulses is about 20 s. The blue streak visible between P and P400s is a sidelobe artefact of direct P, as it remains parallel to P even at short ranges, whereas a converted phase from a discontinuity shallower than 400 km would become increasingly delayed at close ranges. Also visible in Fig. 1c are hints of the discontinuity phases which can be seen more clearly in the vertical-component stack (Fig. 1a).

SH wave stack (transverse component)

Figure 1e shows the result of aligning the SH wave and stacking on the transverse horizontal component at ranges between 40° and 120° and travel times of up to 7 min following the SH wave arrival. The data are binned and averaged in increments of 0.5° in range and 2 s in time. 2,648 seismograms are stacked from

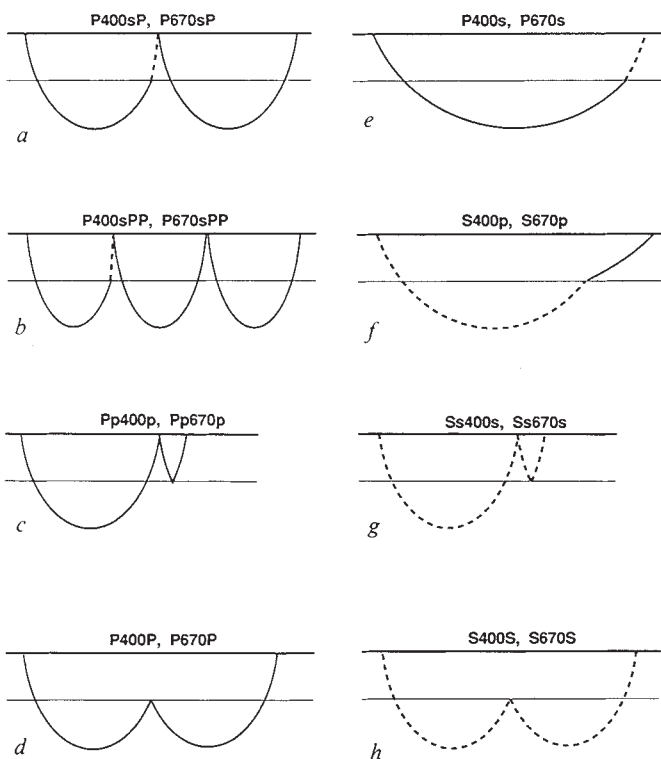


FIG. 3 The secondary seismic phases visible in Fig. 1, which are associated with reflections and conversions at upper-mantle discontinuities. P waves are shown as solid lines and S waves as dashed lines, with the thin horizontal line representing either the 400- or the 670-km discontinuity. Lower-case p and s represent short ray segments between the discontinuities and the surface. Phases a–d can be seen most clearly on the vertical-component stack, e and f most clearly on the radial-component stack, and g and h on the transverse-component stack.

35 stations, for an average of about 17 seismograms per range bin. As in the P wave stacks, the image becomes noisier at ranges beyond about 100° because of the relatively small number of good-quality S_{diff} arrivals in these ranges. Figure 1f shows travel times for seismic phases visible in the transverse-component stack. The S, SS, SSS and ScS phases are seen clearly. As expected⁴³, the SS phase is Hilbert-transformed relative to S.

Also clearly apparent in this stack are the SH wave multiples resulting from topside reflections off the 400-km and 670-km discontinuities (Ss400s and Ss670s; see Fig. 3). These are phase-reversed relative to S and are visible between about 70° and 115°. As in the case of the P wave multiples, a phase also appears between these reflections which is apparently caused by a discontinuity at about 520 km depth. This feature is better resolved in the SH stack because of the relatively large time separation between the Ss400s and Ss670s arrivals. I will discuss later the implications of this phase. Figure 1e also shows travel times for the bottomside reflected phases S400S and S670S, which precede SS. These phases can be seen only faintly in the stacked image. As in the case of the P wave stack, there is no visible phase corresponding to a multiple off a 220-km discontinuity.

Cross-correlation analysis

As an alternative to the stacking procedure discussed above, a cross-correlation method can be used to resolve upper-mantle discontinuity phases. This has the advantage of permitting the computation of discontinuity depths for each seismogram individually to see if these results are consistent with the stacked images. The cross-correlation analysis proceeds as follows: (1) The same seismograms are used as in the stacking method, that is, those from events shallower than 50 km and with signal-to-noise ratios of ten or better. (2) For each reference phase (P or SH), a 41-s wavelet centred on the maximum-amplitude point is used as a reference waveform and the cross-correlation function between this wavelet and the rest of the seismogram is computed. (3) Peaks in the cross-correlation function are identified and saved. This procedure finds the time shifts to pulses in the seismogram that are similar in shape to the reference arrival and which may thus represent other phases. (4) Assuming a particular discontinuity phase of interest and an upper-mantle velocity model, these points may be converted directly into discontinuity depths.

Figure 4 shows the results of this procedure applied to the 2,648 seismograms in the SH wave stack. To resolve the SH wave multiples from upper-mantle discontinuities, points are plotted only if the correlation coefficient, r , between the reference wavelet and the seismogram is less than -0.7 (the correlation coefficient is negative because these arrivals are phase-reversed relative to S). Both Ss400s and Ss670s are visible in this plot, as well as the apparent Ss520s multiple seen in the stacked image. Using an upper-mantle velocity model (such as PREM⁴²), discontinuity depths can be calculated for each point in this plot.

Figure 5 shows the discontinuity depths predicted by cross-correlation analysis, using three different types of discontinuity phases and assuming PREM for the upper-mantle velocities. These phases include P to SV conversions recorded on the radial component between ranges of 40° and 105°, and P wave and SH wave multiples between 70° and 110° recorded on the vertical and transverse components, respectively. The vertical-component P wave is used as the reference wavelet for both the vertical- and radial-component analyses. Points are plotted only for $r > 0.7$ (or $r < -0.7$ for phase-reversed arrivals), and are binned in 20-km depth increments. Calculated average depths to the 400-km discontinuity are 411 km for P400s, 411 km for Pp400p and 415 km for Ss400s. Calculated depths to the 670-km discontinuity are 674 km for P670s, 660 km for Pp670p and 653 km for Ss670s. Depths to the 520-km discontinuity are 520 km and 519 km for Pp520p and Ss520s respectively. These depths depend on the velocity in the upper mantle above the discontinuities: slower velocities imply deeper discontinuities.

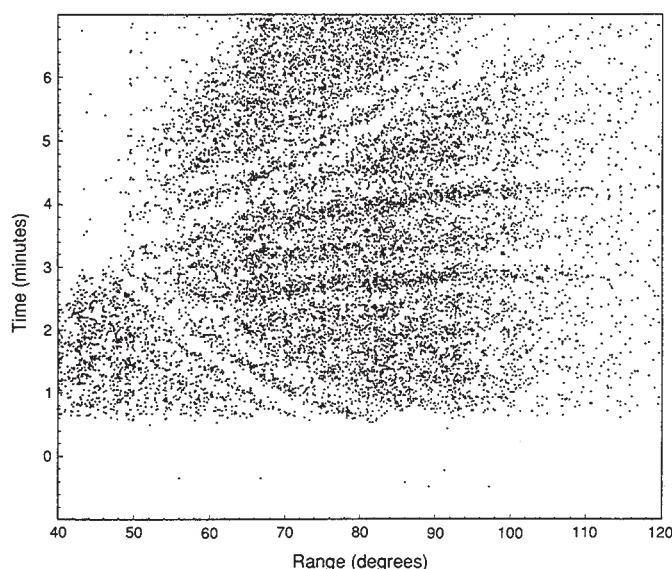


FIG. 4 Cross-correlation peaks for SH wave data (compare with Fig. 1e). The correlation coefficient, r , for each peak is less than -0.7 , in order to resolve the phase-reversed reflections off the top of the upper-mantle discontinuities.

Assuming fixed upper-mantle velocities, uncertainties in these depths are ~ 10 – 20 km; a more formal error analysis is in progress.

As discussed previously, there is no possibility of resolving a 520-km discontinuity in long-period P-to-SV converted arrivals, given the small time separation between the 400-km and 670-km phases. There is no clear evidence for a discontinuity at 220 km depth for any of these converted phases. The P and S wave multiples suggest the possible presence of a discontinuity at ~ 840 km depth. The P wave stack (Fig. 1a) shows a hint of a possible Pp840p phase, but no corresponding S wave multiple can be seen in the SH wave stack (Fig. 1e).

Discussion

These stacks do not show any phases that can be associated with a discontinuity at 220 km depth. This is consistent with many upper-mantle models, which do not have significant structure at this depth^{9,10}. Other models, however, do contain large discontinuities at 220 km, and in some areas the evidence for structure near this depth is quite strong^{5,6}. PREM contains P and S wave impedance contrasts at 220 km that are roughly comparable in size to the contrasts at 400 km. Given that P wave and S wave multiples from the 400-km discontinuity are easily visible in these stacks, a 220-km discontinuity of comparable magnitude should also be seen. The absence of visible 220-km discontinuity phases indicates that the impedance contrast at 220 km must be significantly less than the PREM values or that the 220-km discontinuity is not a globally coherent feature. These results are consistent with recent analysis of ScS multiples which also fails to find evidence for a 220-km discontinuity²⁶.

The fact that 400-km and 670-km discontinuity phases can be seen so clearly in stacked images of global seismic data sets indicates that depths to these discontinuities cannot vary too widely between regions, because otherwise the coherence of the stack would be lost. Results obtained from cross-correlation analyses of records from individual seismic stations generally indicate discontinuity depth variations of less than ± 20 km. Amplitudes of the Ss400s and Ss670s phases in the stacked record section are about 30% smaller than those predicted by the PREM model. This may indicate that the PREM S wave impedance contrast is too large at these discontinuities or it may be a result of slight differences in the discontinuity depths or

lateral velocity variations which cause time shifts and partial cancellation of the pulses during the stacking procedure²⁶.

Both the P wave and S wave stacks suggest the presence of phases resulting from reflections off a 520-km discontinuity, with an apparent amplitude roughly half the size of the 400-km discontinuity phases. However, because these phases occur between the 400-km and 670-km multiples, the possibility must be considered that they are artefacts from the sidelobes of these phases. Detailed synthetic seismogram analysis will be necessary to resolve this question, but several factors suggest that these 520-km phases are real. These include: (1) The 520-km phases are not centred between the 400- and 670-km phases, but are closer in time to the 400-km phases. The time separation is too great for them to be considered sidelobes of the 670-km phases. (2) The 520-km phases follow the 400-km phases, but no corresponding phases follow the 670-km phases. (3) Cross-correlation analyses of the Pp520p and Ss520s phases give the same apparent depth for the 520-km discontinuity despite the different time separation between the 400-, 520- and 670-km multiples, because of the different velocities of P and S. (4) The amplitude of the Ss520s phase is over half the amplitude of the Ss400s phase—too large to be a sidelobe artefact of the 400-km phase, as the positive sidelobes of S are only about 15% of the size of the main S pulse in these stacked records. (5) The waveform of the stacked SS pulse shows no significant change from direct S (after Hilbert-transforming), indicating that the additional crustal transits have had little effect on the waveform. (6) Preliminary analysis of SH synthetic seismograms (toroidal mode sum) computed for PREM indicates no waveforms that could be mistaken for Ss520s.

Most recent upper-mantle velocity models do not have

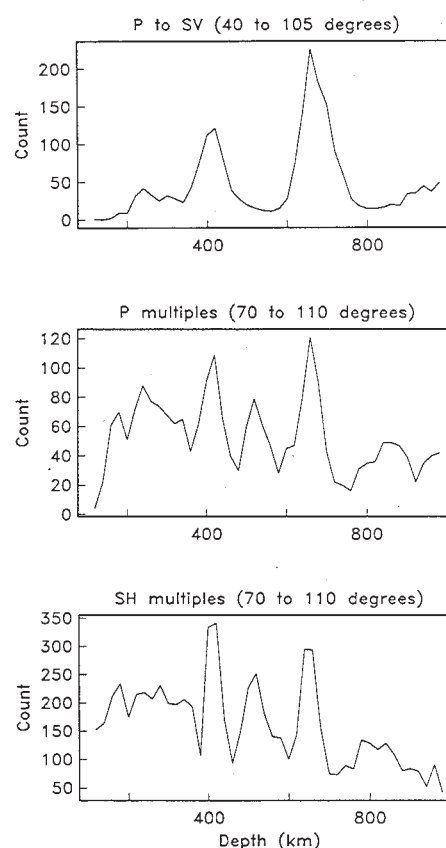


FIG. 5 Number of cross-correlation peaks as a function of discontinuity depths for P-to-SV converted waves (top), P wave multiples (middle), and SH wave multiples (bottom). Results have been binned in increments of 20 km in depth. The P and S multiples suggest the presence of a discontinuity at about 520 km depth.

significant structure near 520 km depth^{9,10,42}, although several older velocity models do contain discontinuities or kinks near this depth^{5,13-17}. These models were based mainly on analyses of P and SH wave refraction data and P wave slowness data recorded at arrays. These types of analyses cannot distinguish between sharp velocity discontinuities and local regions of steep velocity gradients. The long-period data shown in this paper also cannot distinguish between a discontinuity or enhanced velocity gradient at 520 km. The amplitude of the apparent Pp520p and Ss520s phases discussed in this paper are consistent with an impedance contrast at 520 km of ~3%, occurring within a gradient region less than about 25 km thick. But there is also some evidence for a 520-km reflector from quarry blast explosions¹¹ and short-period precursors to the phase P'P' (ref. 12). If these observations do represent reflections of high-frequency energy off a 520-km discontinuity, then the discontinuity must be relatively sharp (less than ~4 km thick), just as other short-period P'P' precursors imply a locally sharp 670-km discontinuity^{44,45}.

It is tempting to explain the possible existence of a 520-km discontinuity as representing the phase change between β -olivine and spinel, which occurs at about the right depth^{1,12}. Recent high-pressure laboratory results, however, suggest that the change in velocity associated with this phase change occurs gradually and is quite small⁴⁶. Other suggestions for possible phase changes near 520 km have included enstatite $\rightarrow$ spinel + stishovite¹² and garnet $\rightarrow$ garnet + perovskite¹, with more recent laboratory results indicating many complicated phase changes in the pyroxene/garnet system⁴⁷. In general, predicted upper-mantle phase changes are highly dependent on composition and temperature, so conclusively identifying a specific phase change may be difficult. □

Received 30 November 1989; accepted 26 January 1990.

1. Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, New York, 1975).
2. Ito, E. & Takahashi, E. *J. geophys. Res.* **94**, 10637-10646 (1989).
3. Anderson, D. L. & Bass, J. D. *Nature* **320**, 321-328 (1986).
4. Lehmann, I. *Ann. Geophys.* **15**, 93-118 (1959).
5. Hales, A. L., Muirhead, K. J. & Rynn, J. M. W. *Tectonophysics* **63**, 309-348 (1980).
6. Drummond, B. J., Muirhead, K. J. & Hales, A. L. *Geophys. J. R. astr. Soc.* **70**, 67-77 (1982).
7. Anderson, D. L. *J. geophys. Res.* **84**, 7555-7560 (1979).
8. Leven, J. H., Jackson, I. & Ringwood, A. E. *Nature* **289**, 234-239 (1981).
9. Walck, M. C. *Geophys. J. R. astr. Soc.* **76**, 697-723 (1984).
10. Grand, S. P. & Helmberger, D. V. *Geophys. J. R. astr. Soc.* **76**, 399-438 (1984).
11. Hoffman, J. P., Berg, J. W. & Cook, K. L. *Bull. seismol. Soc. Am.* **51**, 17-27 (1961).
12. Whitcomb, J. H. & Anderson, D. L. *J. geophys. Res.* **75**, 5713-5728 (1970).
13. Helmberger, D. & Wiggins, R. A. *J. geophys. Res.* **76**, 3229-3245 (1971).
14. Wiggins, R. A. & Helmberger, D. V. *J. geophys. Res.* **78**, 1870-1880 (1973).
15. Helmberger, D. V. & Engen, G. R. *J. geophys. Res.* **79**, 4017-4028 (1974).
16. Simpson, D. W., Mereu, R. F. & King, D. W. *Bull. seismol. Soc. Am.* **64**, 1757-1788 (1974).
17. Fukao, Y. *Geophys. J. R. astr. Soc.* **50**, 621-642 (1977).
18. Byerly, P. *Bull. seismol. Soc. Am.* **16**, 209-265 (1926).
19. Jeffreys, S. H. *Mon. Not. R. astr. Soc. Geophys. Suppl.* **3**, 401-422 (1936).
20. Johnson, L. R. *J. geophys. Res.* **72**, 6309-6325 (1967).
21. Nguyen-Hai *Ann. Geophys.* **15**, 285-346 (1963).
22. Husebye, E. & Madariaga, R. *Bull. seismol. Soc. Am.* **60**, 939-952 (1970).
23. Davies, D., Kelly, E. J. & Filson, J. R. *Nature* **232**, 8-13 (1971).
24. Gutowski, R. R. & Kaneshewich, E. R. *Geophys. J. R. astr. Soc.* **36**, 21-32 (1974).

25. Ward, S. N. *Bull. seismol. Soc. Am.* **68**, 133-153 (1978).
26. Revenaugh, J. & Jordan, T. H. *J. geophys. Res.* **94**, 5787-5813 (1989).
27. Engdahl, E. R. & Flinn, E. A. *Science* **163**, 177-179 (1969).
28. Bolt, B. A. *Geophys. J. R. astr. Soc.* **20**, 367-382 (1970).
29. Adams, R. D. *Bull. seismol. Soc. Am.* **61**, 1441-1451 (1971).
30. Husebye, E. S., Haddon, R. A. W. & King, D. W. *J. Geophys.* **43**, 535-543 (1977).
31. Nakanishi, I. *Geophys. J.* **93**, 335-346 (1988).
32. Davis, J. P., Kind, R. & Sacks, I. S. *Geophys. J.* **99**, 595-604 (1989).
33. Vinnik, L. P. *Phys. Earth planet. Inter.* **15**, 39-45 (1977).
34. Faber, S. & Müller, G. *Bull. seismol. Soc. Am.* **70**, 487-508 (1980).
35. Barley, B. J., Hudson, J. A. & Douglas, A. *Geophys. J. R. astr. Soc.* **69**, 159-172 (1982).
36. Bock, G. & Ha, J. *Geophys. J. R. astr. Soc.* **77**, 593-615 (1984).
37. Faber, S. & Müller, G. *J. Geophys.* **54**, 183-194 (1984).
38. Baumgardt, D. R. & Alexander, S. S. *Bull. seismol. Soc. Am.* **74**, 1683-1702 (1984).
39. Kind, R. & Vinnik, L. P. *J. Geophys.* **62**, 138-147 (1988).
40. Paulssen, H. *J. geophys. Res.* **93**, 10489-10500 (1988).
41. Bock, G. *J. Geophys.* **64**, 73-81 (1988).
42. Dziewonski, A. M. & Anderson, D. L. *Phys. Earth planet. Inter.* **25**, 297-356 (1981).
43. Choy, G. L. & Richards, P. G. *Bull. seismol. Soc. Am.* **65**, 55-70 (1975).
44. Richards, P. G. *J. Geophys.* **38**, 517-527 (1972).
45. Lees, A. C., Bukowski, M. S. T. & Jeanloz, R. *J. geophys. Res.* **88**, 8145-8159 (1983).
46. Weidner, D. J., Sawamoto, H., Sasaki, S. & Kumazawa, M. *J. geophys. Res.* **89**, 7852-7860 (1984).
47. Jeanloz, R. & Thompson, A. B. *Rev. Geophys. Space Phys.* **21**, 51-74 (1983).

ACKNOWLEDGEMENTS. I thank J. Garmany, G. Masters and G. Nolet for advice regarding this work. This research was partially supported by the NSF.

***In vivo* alteration of telomere sequences and senescence caused by mutated *Tetrahymena* telomerase RNAs**

Guo-Liang Yu, John D. Bradley, Laura D. Attardi & Elizabeth H. Blackburn*

Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA

Mutating the CAACCCCAA sequence in the RNA component of telomerase causes the synthesis *in vivo* of new telomere sequences corresponding to the mutated RNA sequence, demonstrating that the telomerase contains the template for telomere synthesis. These mutations also lead to nuclear and cell division defects, and senescence, establishing an essential role for telomerase *in vivo*.

TELOMERES, the ends of eukaryotic chromosomes, are essential for the stabilization of linear chromosomes and the completion of chromosomal DNA replication^{1,2}. Telomeres from diverse

eukaryotes, including protozoans, yeasts, plants and mammals, consist of simple tandemly repeated G+C-rich sequences, with the G-rich sequence at the 3' end of each DNA strand of a chromosome²⁻⁵. For example, the telomeric sequence of the ciliated protozoan *Tetrahymena thermophila* consists of TTGGGG-AACCCC repeats⁶.

The G-rich strand of telomeric DNA is synthesized *in vitro* by telomerase⁷, a ribonucleoprotein enzyme with an essential RNA component^{8,9}. The *Tetrahymena* telomerase synthesizes repeats of the telomeric DNA sequence 5'-TTGGGG-3' without requiring an added exogenous template⁷⁻⁹. Analogous telomerase activities have been identified in the ciliates *Oxytricha nova*¹⁰ and *Euplotes crassus*¹¹, and recently in human (HeLa) cells¹²; each synthesizes its species-specific repeats onto the 3' end of a telomeric sequence using the G-rich sequence as a primer.

Telomerase activity, originally proposed on the basis of *in vivo* studies of heterologous telomeres in yeast, readily explains

*To whom correspondence should be addressed.

several aspects of telomere behaviour *in vivo*¹³⁻¹⁹. Telomeres with various sequences are recognized as such and function as telomeres when transformed into yeast^{13,20-22}; yeast-specific repeats are invariably added to the distal end of the foreign telomeric sequence^{13,20}. This behaviour parallels the recognition of different telomeric sequences as primers by telomerases *in vitro*⁷⁻¹². In many species the number of telomeric repeats on the end of a particular chromosome is variable, resulting in length heterogeneity of telomeric restriction fragments and their appearance as diffuse bands on agarose electrophoresis gels. This length heterogeneity could be caused by a dynamic equilibrium between incomplete semiconservative chromosome replication and telomere synthesis by telomerase^{2,16}, but the role of telomerase in telomere maintenance *in vivo* has not been tested directly.

The telomerase RNA genes from *T. thermophila* and *E. crassus* have been identified, cloned and sequenced^{9,23}. Strikingly, both of these RNAs include a sequence that could serve as a template for the synthesis of species-specific telomeric repeats. The integrity of this sequence—5'-CAACCCCAA-3'—in the *Tetrahymena*

mena telomerase RNA is essential for telomerase activity⁹. Furthermore, the analogous sequence 5'-CAAAACCCCAA-3' in the *Euplotes* telomerase RNA has been identified as the template strand sequence for *in vitro* synthesis of the repeated *Euplotes* telomeric sequence, TTTTGGGG (ref. 23). Although these *in vitro* experiments implicate the telomere-complementary sequence in the RNA as the template for telomere synthesis *in vitro*, there has been no direct demonstration that they function as templates *in vivo*.

We now present direct genetic evidence that the telomerase RNA contains the template for telomere synthesis *in vivo*. We altered the cloned *Tetrahymena* telomerase RNA gene in its putative template region by site-directed mutagenesis. Using a high-copy-number vector, we transformed *Tetrahymena* cells with three different mutant telomerase RNA genes. Two of these resulted in the *in vivo* synthesis of new telomeric sequences corresponding to the mutation in the template. In addition, telomeres in these transformants were longer than control cell telomeres. The third mutation led to telomere shortening. All three mutations, however, caused unexpected and striking phenotypic changes: specifically, nuclear and cell division were impaired, and viability was greatly decreased or the cells became senescent. These findings uncover an apparently new role for telomeres *in vivo* and indicate that the telomerase RNA gene is an essential gene.

Telomerase RNA overexpression

We subcloned the wild-type telomerase RNA gene (termed here the TER gene) from *T. thermophila*⁹ into a circular ribosomal DNA plasmid vector, prD4-1. This vector, described in detail in ref. 24, carries the *cis*-acting replication control elements of the amplified rDNA and therefore replicates at high copy number (> 10³ copies per cell) in the macronucleus of *T. thermophila*. The polygenomic macronucleus, derived from the germ-line micronucleus, is the primary site of gene expression in ciliates. We inserted a 550-base pair (bp) *Hind*III-*Dde*I fragment containing the coding sequence for the 159-nucleotide *T. thermophila* telomerase RNA and its flanking regions⁹ into the *Xho*I site in the polylinker in the plasmid vector to form plasmid prTER1_{wt}. This construct contains a ribosomal RNA gene carrying the dominant selectable paromomycin resistance marker (*Pmr1*) from the plasmid prD4-1 vector, and the wild-type telomerase RNA gene. The telomerase RNA transcription unit is in the same orientation as the ribosomal RNA gene (Fig. 1).

We transformed *T. thermophila* cells with circular plasmid prTER1_{wt} by microinjection into the macronucleus as previously described^{25,26}. Because the vector has a replication advantage over the recipient rDNA²⁴, the injected plasmid was maintained at high copy number. After growing the cells in nonselective medium for three days, we identified transformants by their ability to grow in selective medium containing 100 µg ml⁻¹ paromomycin.

We obtained two prTER1_{wt} transformants from 12 viable microinjected cells in one experiment, and three more independent transformants from pools of microinjected cells from another experiment. Southern-blot analysis showed that the telomerase RNA gene in these transformants, which is normally present as a single copy per haploid genome in the polygenomic macronucleus⁹, was maintained on the amplified rDNA vector for more than 50 generations after microinjection (ref. 24, and data not shown).

To investigate the transcription of the introduced telomerase RNA gene, we isolated macronuclei from equal numbers of cells from a prTER1_{wt} transformant line and from a control line transformed with the prD4-1 vector alone, and assayed them in nuclear run-off experiments. In prTER1_{wt} transformants the telomerase RNA gene was transcribed at significantly higher levels than it was in the control prD4-1 transformant (Fig. 2a). To confirm that the increased rate of transcription reflected expression of the introduced telomerase RNA gene, we analysed

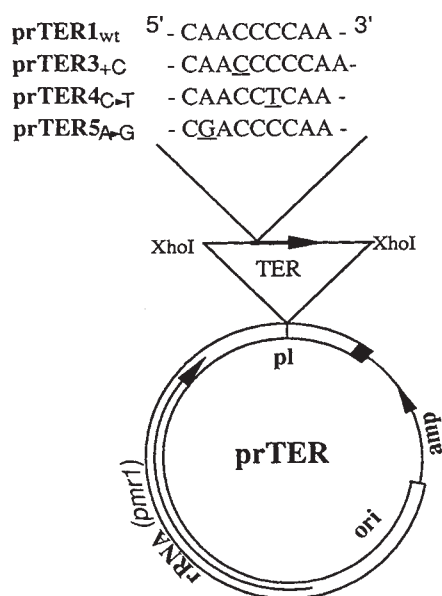


FIG. 1 Construction of prTER plasmids. The vector is shown as a circle. Thick open bar, rDNA sequences; thick solid bar, telomeric sequence; thin line, pBR322 sequences. The arrows show the transcription units of rRNA genes carrying the paromomycin resistance marker (*Pmr1*), ampicillin resistance gene (*amp*) and the telomerase RNA gene (*TER*). The *TER* genes were inserted in the polylinker (*pl*). The putative template sequences (+43 to +51 in the wild-type coding sequence)⁹ are shown, with the mutated positions underlined. The telomerase RNA gene is transcribed in the same direction as the ribosomal RNA gene in these plasmids.

METHODS. A 550-bp *Hind*III-*Dde*I fragment containing the telomerase RNA gene was isolated from clone pCG1 (ref. 9). The fragment was blunt-ended, and *Xho*I linkers were ligated on both ends of the fragment. After digestion with *Xho*I, the fragment was inserted into the polylinker of prD4-1 to construct prTER1_{wt}. Site-directed mutagenesis was accomplished by a variation of the method of Kunkel⁴⁶. A 1.6-kilobase (kb) *Hind*III fragment containing the *TER* gene and some of the flanking rDNA sequence from prTER1_{wt} was inserted into the *Hind*III site of pUC119, and the plasmid was transformed into *Escherichia coli* strain RZ1032. Co-infection with bacteriophage M13K07 resulted in the production of single-stranded uracil-containing plasmid. Oligonucleotides containing the desired mutations were annealed to the plasmid and used as primers for synthesis of the second strand. The DNA was transformed into *E. coli* strain DH5, and mutants were identified by sequencing single-stranded DNA obtained from individual transformants. In each of the three cases, ~50% of transformants carried the properly mutated plasmid. The 550-bp *Xho*I fragments containing the mutagenized *TER* genes were subcloned into the polylinker of the vector prD4-1.

RNA from cells transformed with a vector carrying a mutant telomerase RNA gene, prTER3₊C (described below). Because this mutant *TER* gene contains an additional C residue in the putative template region (see below, and Fig. 1), we expected its correct transcript to be one nucleotide longer than the wild-type transcript. Indeed, in primer extension experiments using total cellular RNA isolated from a prTER3₊C transformant line, the main primer extension product was one nucleotide longer than that for the control RNA from the untransformed recipient cell strain (Fig. 2b). We confirmed this with an additional transformant line (data not shown). Therefore, the telomerase RNA in prTER3₊C transformants contained the correct 5' end, and most of the telomerase RNA in these cells was of the mutant

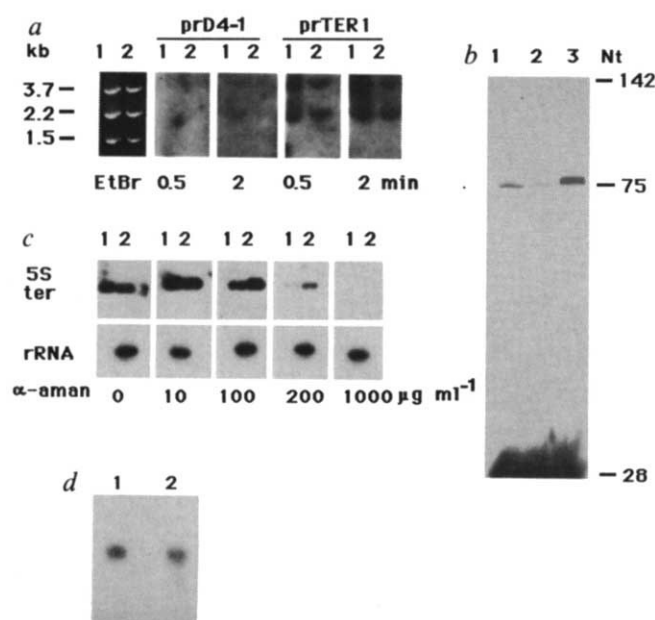


FIG. 2 Overexpression of the telomerase RNA gene. *a*, Nuclear run-off experiments. An equal number of nuclei isolated from prTER1_{wt} and control prD4-1 transformant cells in late log phase ($\sim 2 \times 10^5$ cells ml⁻¹) were incubated with [α -³²P] rUTP and cold rATP, rCTP and rGTP in a buffer containing 0.1 M KCl, 20 mM HEPES-KOH buffer (pH 7.9), 5 mM dithiothreitol, 5% glycerol, 0.1 mM EDTA, 6 mM MgCl₂ (ref. 47) for 0.5 or 2 min at 30 °C. After treatment with protease K and phenol extraction, the radiolabelled RNA was used to probe Nytran filters carrying the immobilized cloned telomerase RNA gene. As shown in the ethidium bromide (EtBr)-stained photograph, each lane contains a mixture of a 3.7-kb uncut plasmid carrying the telomerase RNA gene and two restriction fragments (2.2 and 1.5 kb) resulting from digestion with *Bgl*II and *Xmn*I. *Bgl*II cuts at the beginning of the telomerase RNA coding region to generate the 2.2-kb fragment containing almost all the coding region, and the 1.5-kb fragment containing the upstream sequences. Lanes 1 and 2 contained 1 μ g and 0.3 μ g DNA, respectively. *b*, Primer extension of RNA from a prTER3₊C transformant. Total RNA was isolated using guanidinium isothiocyanate⁴⁸. To map the 5' end of the telomerase RNA in prTER3₊C transformants, a 28-nucleotide oligonucleotide complementary to the coding sequence (+47 to +64), downstream from the additional C, was labelled by polynucleotide kinase and extended using reverse transcriptase. Lane 1, recipient; lane 2, prTER3₊C transformant from which the telomerase RNA gene had been lost (data not shown); lane 3, prTER3₊C transformant containing the mutant telomerase RNA gene. Nt, nucleotides. *c*, Nuclear run-off assays in the presence of α -amanitin (α -aman). Assays were performed as described in *a*. Top panel, prTER1_{wt} (lane 1) and *Tetrahymena* 5S DNA (lane 2); bottom panel, prD4-1 (rDNA probe). Concentrations of α -amanitin are shown at the bottom of the panel. *d*, Northern blot analysis of TER RNA. Lane 1, recipient; lane 2, prTER1_{wt} transformant. RNA was fractionated by agarose gel electrophoresis and the blot probed with nick-translated TER gene.

type. Furthermore, the sensitivity of telomerase RNA gene transcription to α -amanitin in nuclear run-off experiments was similar to that of 5S RNA, rather than to that of rRNA (Fig. 2c), indicating that most of the labelled telomerase RNA in prTER1_{wt} transformants was not a read-through product of RNA polymerase I transcription, but was expressed from its own promoter. Despite the increased rate of transcription resulting from its high gene dosage, northern-blot analysis showed that the accumulated level of telomerase RNA was the same in prTER1_{wt} transformants as in control cells (Fig. 2d). Because telomerase is a complex of protein as well as RNA components^{8,9}, it is likely that the excess telomerase RNA in the transformants is degraded if it is unable to assemble into the telomerase ribonucleoprotein.

We found no significant difference in growth rates or morphology between cells of prTER1_{wt} transformant lines and cells transformed with prD4-1 vector alone. In addition, telomere lengths in these prTER1_{wt} transformants (examined in five independent prTER1_{wt} transformant lines and in 14 single cell subclones from one of these lines) were the same as in untransformed controls (see below). Southern blotting showed that all these prTER1_{wt} transformants had retained the introduced telomerase RNA gene at high copy numbers (data not shown). The absence of phenotypic changes in these wild-type gene transformants was consistent with the observed normal levels of accumulated telomerase RNA.

Altered telomere lengths

Using oligonucleotide-directed mutagenesis, we made three separate single nucleotide mutations in the telomere-complementary sequence of the telomerase RNA gene. These mutant plasmids were termed prTER3₊C, prTER4_{C→T} and prTER5_{A→G}, the nucleotide alteration being indicated as a subscript (Fig. 1). We microinjected each mutant telomerase RNA gene in the rDNA vector into *Tetrahymena* macronuclei.

We analysed telomere length by Southern-blot analysis of *Bam*HI-digested DNA; as a probe we used [³²P] 5' end-labelled telomeric (T₂G₄)₂ (top panel in Fig. 3), because we reasoned that the original telomeric sequence was likely to be still present in all telomeres, even if additional new telomeric sequences were synthesized. After *Bam*HI digestion, the high copy number rDNA telomeric fragment (shown in Fig. 3) was well separated from other macronuclear telomeric fragments, most of which migrated near limit mobility in agarose gel electrophoresis. The telomeric rDNA fragments in prTER3₊C and prTER5_{A→G} transformants (lanes 3 and 5) had longer mean lengths and broader size distributions than did the telomeric fragments in control recipient cells and prTER1_{wt} transformants (lanes 1 and 2). The telomeric fragments of other macronuclear chromosomes also appeared to be correspondingly longer and broader (data not shown). By contrast, in the prTER4_{C→T} transformant (lane 4) the rDNA telomeric fragment was significantly shorter than it was in the recipient strain and prTER1_{wt} transformants.

The long telomere phenotype depended on the presence of the mutant telomerase RNA gene. We analysed telomere lengths in single cell subclones of a prTER3₊C transformant by Southern blotting and probing with an rDNA fragment that hybridized to the telomeric fragment (Fig. 4). All of these subclones were derived from a single transformant clone that initially contained the mutant telomerase RNA gene at high copy number (data not shown) and long telomeres (Fig. 4a, lane 3). Of the 15 subclones that grew sufficiently for DNA analysis, seven still contained high, although variable, copy numbers of the mutant telomerase RNA gene on rDNA replicons; all of these seven contained long telomeres (Fig. 4b, lanes 1-3, 5, 11, 13 and 15). The remaining eight subclones had lost the mutant telomerase RNA gene through recombination of the rDNA vector with endogenous rDNA (refs 26, 27), and contained telomeres of normal length (Fig. 4b, lanes 4, 6-10, 12 and 14). That telomere length reverted to normal within 30 cell generations after the

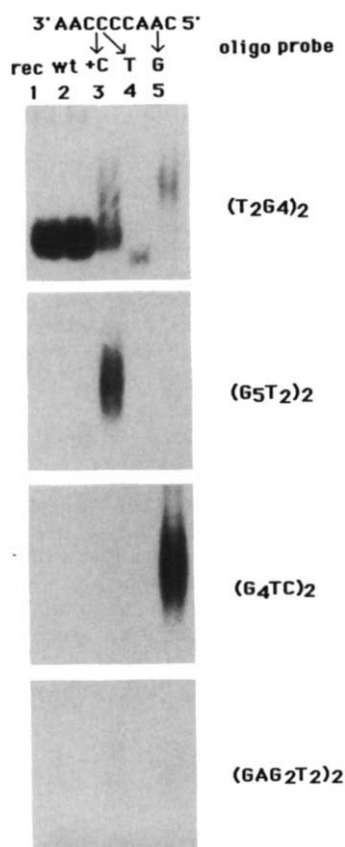


FIG. 3 Cells transformed by mutant telomerase RNA genes contain altered telomeric sequences. Genomic DNA from recipient (rec, lane 1), and cells transformed by prTER1_{wt} (lane 2), prTER3_{+C} (lane 3), prTER4_{C→T} (lane 4) and prTER5_{A→G} (lane 5), were digested with *Bam*HI and four identical sets of samples fractionated by agarose electrophoresis; blots were probed with oligonucleotides (oligo) shown on the right. wt, Wild type.

telomerase RNA gene was lost indicates that the presence of the mutant telomerase RNA is necessary for maintenance of long telomeres.

Novel telomere sequences

To analyse the telomeric sequences in the transformants, we probed identical blots containing equal aliquots of the set of transformant and control DNA samples with oligonucleotides with sequences corresponding to each mutation in the

telomerase RNA gene (see Fig. 1): (G₅T₂)₂, (GAG₂T₂)₂ and (G₄TC)₂. Oligonucleotide (G₅T₂)₂ hybridized only to the telomeric fragments in the prTER3_{+C} transformant (Fig. 3, lane 3). Conversely, the (G₄TC)₂ probe hybridized only with the telomeric fragment in the prTER5_{A→G} transformant (lane 5). By contrast, we could not detect any hybridization of the (GAG₂T₂)₂ probe that was specific to telomeric fragments in the prTER4_{C→T} transformant (lane 4), even though we used various hybridization and post-hybridization wash conditions and autoradiographic exposures. We conclude that in the prTER4_{C→T} transformant, unlike in the other two mutant telomerase RNA gene transformants, telomeres do not contain the altered telomeric sequence predicted from the mutation in the telomerase RNA.

To confirm that the mutated telomerase RNA gene dictates the synthesis of the predicted telomeric sequence *in vivo*, we cloned a telomeric fragment from a prTER5_{A→G} transformant. The insert in one clone that hybridized to (G₄TC)₂ was identified as an rDNA telomere by Southern blotting and DNA sequence analysis. The telomeric sequence of this clone is shown in Fig. 5. It includes several tandem repeats of the mutant telomeric sequence G₄TC located distally to a run of 30 G₄T₂ repeats. It is interesting that at two positions, a single wild-type repeat interrupts the total of 11 mutant repeat units, and an additional stretch of distal wild-type G₄T₂ repeats is also present.

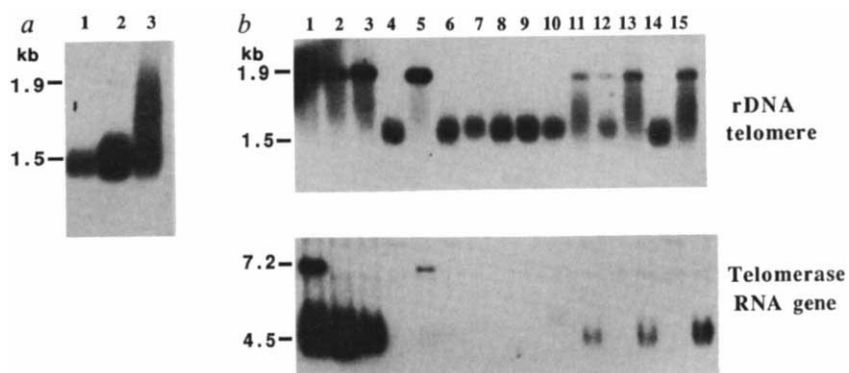
The telomeric sequence is normally highly invariant in *Tetrahymena*; in previous studies, sequencing of > 2,400 nucleotides in 14 different telomere clones from wild-type *Tetrahymena* revealed only one different repeat—G₅T₂ (refs 28, 29). This contrasts with the variability that is characteristic of telomeric repeat sequences in some organisms (ref. 4; see ref. 3 for a review). We conclude that the mutant telomere containing G₄TC repeats was synthesized by copying the altered sequence in the mutated telomerase RNA.

Mutant phenotypes

Each of the three mutated telomerase RNA genes tested caused striking morphological changes and senescence when overexpressed in *Tetrahymena* cells. The prTER4_{C→T} transformants showed the most extreme senescence phenotype. We analysed three transformant lines, and isolated 48 and 24 single cells from two of them. None reverted and, after a week in culture, all 72 sublines were dead. Before they died, most prTER4_{C→T} transformant cells had rounded or irregular shapes, and were much larger than wild-type cells (Fig. 6, compare *a* and *c*). The macronuclei were also enlarged. The macronuclei were often highly irregular in shape and stained variably with (4',6'-diamidino-2-phenylindole (DAPI); Fig. 6*c*).

Using plasmid prTER3_{+C}, we obtained nine independent transformant lines. We also observed striking morphological and divisional changes in these lines. The prTER3_{+C}

FIG. 4 Presence of a mutant telomerase RNA gene is required for long telomeres. *a*, Ribosomal DNA telomeres in recipient (lane 1), prTER1_{wt} transformant (lane 2), and a prTER3_{+C} transformant (lane 3). *b*, Southern blotting analysis of rDNA telomeres in single cell lines. Single cell lines from a prTER3_{+C} transformant were isolated into 1 ml nonselective medium in multi-well tissue culture plates. Southern blots of DNA samples from 15 single cell lines were probed with rDNA or TER gene probes. Top panel, rDNA telomeres; bottom panel, TER gene. Previous work^{24,26,27} showed that the rDNA vector recombines with the endogenous rDNA to form a linear molecule. The 7.2-kb fragment contains the TER gene in the circular form, whereas the 4.5-kb telomeric fragment contains the TER gene in the linear recombinant form.



transformant cells grew extremely large, with very irregular cell and nuclear shapes. Staining with DAPI (Fig. 6b) indicated that although macronuclear and cell divisions were greatly impaired, DNA replication was relatively unaffected. Multiple small nuclei, tentatively identified as micronuclei, were frequently visible (Fig. 6b). It is interesting that these phenotypic features are similar to those of the 'monster cells' generated at nonpermissive temperatures in temperature-sensitive cell division arrest (*cda*) mutants, isolated previously in *T. thermophila*^{30,31}. The molecular basis of the *cda* mutants is not known.

To examine the senescence phenotypes of the prTER3_{+C} transformants, we isolated, 3 days after the addition of paromomycin, each of 48 drug-resistant single cells from one transformant line into 1 ml nonselective medium. All subclones had monster phenotypes and grew very slowly, with 25 dying and only 23 growing sufficiently for DNA analysis. All 23 subclones had long telomeres (data not shown). From an 11-day-old stock culture of another transformant, we made 24 single-cell subclones. Nine subclones failed to divide further, whereas eight grew to saturation after one week, at which time most of the cells appeared normal. These eight saturated cultures were overgrown by cells that had lost the introduced telomerase RNA gene (see above and Fig. 4). But the remaining seven subclones grew only very slowly after one week, had monster phenotypes, and subsequently all single cells isolated from them failed to divide further—that is, they were senescent. These seven subclones had retained the mutant telomerase RNA gene and had

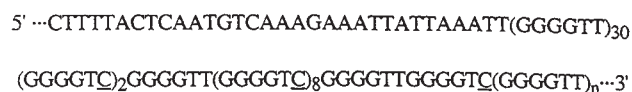


FIG. 5 Sequence of a novel telomere from transformed *Tetrahymena*. Total genomic DNA from a prTER5_{A→G} transformant was treated with T4 DNA polymerase in the presence of all four deoxynucleoside triphosphates (1 mM) to blunt the telomeric ends of the chromosomal DNA. The DNA sample was then digested with *Hind*III and ligated into the polylinker in plasmid pUC119 that had been linearized by digesting with *Hind*III and *Sma*I. The ligated DNA was transformed into *E. coli* and plated on LB plates containing 50 µg ml⁻¹ ampicillin. Clones containing telomeric sequences were screened by colony hybridization, using labelled (G₄TC)₂ oligonucleotide as a probe. Southern analysis identified the cloned sequence as an rDNA telomere. The telomeric sequence of this clone is shown.

long telomeres (see above and Fig. 4). We observed similar results to those seen with the prTER3_{+C} transformant with the three prTER5_{A→G} transformant lines that we obtained; the 'monster' phenotype of these lines is shown in Fig. 6d.

In summary, all three alterations of the telomere-complementary sequence in the telomerase RNA caused impairment and eventual arrest of cell division in *Tetrahymena*; reversion of this phenotype occurred upon loss of the mutant telomerase RNA gene.

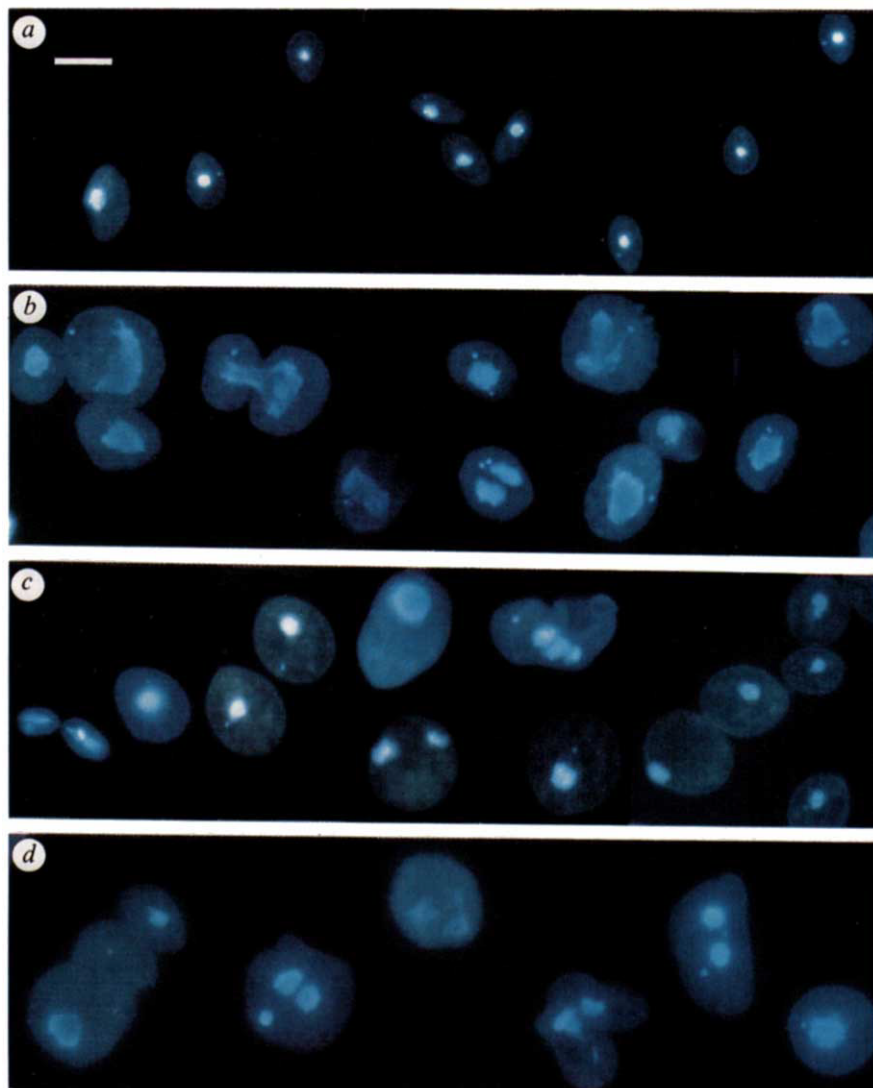


FIG. 6 *Tetrahymena* cells transformed with mutant telomerase RNA genes. Cells were fixed in 3.7% paraformaldehyde and stained with 0.1 µg ml⁻¹ DAPI to visualize nuclei and show the shape of the cells, and photographed using a Zeiss Universal microscope. a, Wild-type cells; b, plasmid prTER3_{+C} transformants; c, plasmid prTER4_{C→T} transformants; d, plasmid prTER5_{A→G} transformants. Bar, 50 µm.

Discussion

Telomerase RNA template. We changed the sequences of telomeres *in vivo* by transformation with two different mutated telomerase RNA genes containing an altered template sequence. Specifically, overexpression of the mutant telomerase RNA gene containing the sequence 5'-CGACCCCAA-3' (mutated base underlined) in transformed cells resulted in incorporation of 5'-GGGGTC-3' repeats. The presence of more-distal wild-type repeats in this mutant telomere suggests that although the mutant repeats were added onto the wild-type telomeric GGGGTT repeats, recombination between telomeres or growth after the mutant gene was lost from this cell had occurred. We similarly found that the mutated sequence 5'-CAACCCCAA-3' also specifies the synthesis of 5'-GGGGGT-3' repeats. These results establish that telomerase uses the telomeric sequence in the RNA as a template for telomere synthesis *in vivo*.

Telomere length regulation. The telomeres from the prTER3_C and prTER5_{A→G} transformants, which had altered telomere sequences, were longer than those in control cells. This indicates that the dynamic equilibrium between elongation and shortening of telomeres *in vivo* was shifted towards elongation. Although it is possible that these two changes in the telomerase template increase its activity *in vivo*, other data suggest that *in vivo*, a negative regulator of telomerase activity competes with telomerase for binding to the end of telomeres³². If changing the terminal sequences of a telomere reduces its binding affinity for the negative regulator, net telomerase activity could effectively increase. A candidate for such a negative regulator is a tightly binding sequence-specific telomere protein of the class that has been characterized in the ciliate *Oxytricha*³³⁻³⁷. The *Oxytricha* telomere protein has greater affinity for the *Oxytricha* telomeric sequence than for the *Tetrahymena* telomeric sequence³³. When bound by these structural telomere proteins *in vitro*, chromosomal DNA ends are protected from exonuclease and methylation³³⁻³⁷. The rDNA telomeres in *Tetrahymena* are in a non-nucleosomal complex^{38,39}, which is likely to include analogous tightly bound proteins (E. H. Henderson, personal communication).

The prTER4_{C→T} transformants did not contain the altered telomeric sequence predicted from the change in the telomerase RNA template, and telomere length was decreased from ~300 to <200 bp of telomeric repeats compared with controls. Although other explanations are not ruled out, the results with prTER4_{C→T} transformants strongly indicate that the C→U change in the telomerase in these cells interferes with its activity, and that impairing the addition of telomeric repeats leads to telomere shortening and eventual loss, which would explain why the transformants become senescent. The inactivation could simply reflect inefficient use of dATP by the enzyme, because copying of the mutated template sequence would require incorporation of dATP by telomerase. Alternatively, the mutated C nucleotide could have some other function in the activity or formation of the telomerase complex. By overexpressing other mutated telomerase RNA genes we hope to test whether functions can be assigned to this nucleotide, as well as other specific nucleotides or regions in the telomerase RNA.

The telomerase RNA gene is the first non-ribosomal RNA gene to be used to transform *Tetrahymena* macronuclei. Although the endogenous wild-type gene was still present, the excess product from the introduced mutated gene produced a 'dominant' phenotype. The ability to overexpress genes in this ciliate system should be useful for other studies.

A novel function for telomeres. The telomere shortening and senescence phenotypes of the prTER4_{C→T} transformants strikingly parallel those of yeast cells that lack a functional *EST1* gene¹⁷, and are consistent with the suggestion^{17,40} that this gene encodes a component of telomerase. The unexpected observation that transformants with either lengthened or shortened telomeres had impaired cell division and senescence phenotypes now implicates telomeres in a previously unrecognized role. The

increased length of the telomeres in prTER3_C and prTER5_{A→G} transformants was unlikely to have been deleterious, because wild-type telomeres of similar lengths are not associated with any growth or division abnormalities¹⁵. Likewise, the existence of certain heterologous telomeric sequences in yeast transformants is not associated with a discernible phenotype^{13,20}. But when these heterologous telomeric sequences are maintained, the most distal sequences are invariably yeast-specific. By contrast, in the study described here, we expect that the novel telomeric sequence was added and present, at least initially, at the distal end of the normal repeat sequences. This indicates that the change in the distal telomeric sequence causes the observed deleterious effects in *Tetrahymena*.

That similar morphological and senescence phenotypes were observed with all three mutated genes suggests that they have a common basis. This is most likely to be an impaired interaction with a sequence-specific telomere binding protein analogous to that characterized in *Oxytricha*³³⁻³⁷. Our results indicate the involvement of such a sequence-specific telomeric DNA-protein interaction in cell or nuclear division. We propose that these protein-DNA interactions are either weakened in the transformants with altered telomere sequences, or eventually lost in the transformants with shortened telomeres, with similar consequences for both. In prTER4_{C→T} transformants, loss of cell viability from eventual macronuclear chromosome loss could be superimposed on this effect. This would explain the greater severity of the senescence phenotype in these cells.

How could telomeres be involved in nuclear and cell division? Telomeres are often found in the lamina under the nuclear envelope⁴¹. The telomeric DNA-protein complex of *Oxytricha* self-associates into very large polymeric forms³⁷, and *in vitro*, synthetic telomeric DNA sequences are bound by the intermediate filament proteins vimentin and nuclear lamins^{42,43}. These observations indicate a possible association of telomeres with nuclear lamins. Second, macronuclear division in *Tetrahymena* and other ciliates does not involve nuclear envelope breakdown or visible centriolar organization of a mitotic spindle, and segregation of the acentric linear chromosomes that comprise the macronuclear genome depends on mechanisms not involving centromeres. We note that in yeast, segregation of 2-micron circles and HM silencer-containing acentric plasmids requires, respectively, the proteins REP1 and SIR4 (ref. 44), which both share amino-acid sequence similarities to conserved domains in nuclear lamins⁴⁵. Taken together, these observations suggest that telomeric interaction with nuclear lamin-like proteins could be involved in the acentric chromosome segregation and nuclear division of the macronucleus. Further analysis will be necessary to dissect the mechanism of this functional involvement of telomeric DNA structure in nuclear division and cell viability. □

Received 22 December 1989; accepted 23 January 1990.

- Blackburn, E. H. *Cell* **37**, 7-8 (1984).
- Blackburn, E. H. & Szostak, J. W. *A. Rev. Biochem.* **53**, 163-194 (1984).
- Forney, J., Henderson, E. R. & Blackburn, E. H. *Nucleic Acids Res.* **15**, 9143-9152 (1987).
- Richards, E. R. & Ausubel, F. M. *Cell* **53**, 127-136 (1988).
- Moyzis, R. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6622-6626 (1988).
- Blackburn, E. H. & Gall, J. G. *J. molec. Biol.* **120**, 33-53 (1978).
- Greider, C. W. & Blackburn, E. H. *Cell* **43**, 405-413 (1985).
- Greider, C. W. & Blackburn, E. H. *Cell* **51**, 887-898 (1987).
- Greider, C. W. & Blackburn, E. H. *Nature* **337**, 331-337 (1989).
- Zahler, A. M. & Prescott, D. M. *Nucleic Acids Res.* **16**, 6953-6972 (1988).
- Shippen-Lentz, D. & Blackburn, E. H. *Molec. cell. Biol.* **9**, 2761-2764 (1989).
- Morin, G. B. *Cell* **59**, 521-529 (1989).
- Shampay, J., Szostak, J. W. & Blackburn, E. H. *Nature* **310**, 154-157 (1984).
- Bernards, A., Michels, P. A. M., Lincke, C. R. & Borst, P. *Nature* **303**, 592-597 (1983).
- Larson, D. D., Spangler, E. A. & Blackburn, E. H. *Cell* **50**, 477-483 (1987).
- Shampay, J. & Blackburn, E. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 534-538 (1988).
- Lundblad, V. & Szostak, J. W. *Cell* **57**, 633-643 (1989).
- Sugawara, N. & Szostak, J. W. *Yeast* **2** Suppl. 373 (1986).
- Perrot, M., Barreau, C. & Begueret, J. *Molec. cell. Biol.* **7**, 1725-1730 (1987).
- Pluta, A. F. & Zakian, V. A. *Nature* **337**, 429-433 (1989).
- Cross, S. H. et al. *Nature* **338**, 771-774 (1989).
- Brown, W. R. A. *Nature* **338**, 774-776 (1989).
- Shippen-Lentz, D. & Blackburn, E. H. *Science* **247**, 546-552 (1990).
- Yu, G.-L. & Blackburn, E. H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8487-8491 (1989).

25. Tondravi, M. M. & Yao, M.-C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4369–4373 (1986).
26. Yu, G.-L., Hasson, M. & Blackburn, E. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5151–5155 (1988).
27. Yu, G.-L. & Blackburn, E. H. *Molec. cell. Biol.* (in the press).
28. Budarf, M. L. & Blackburn, E. H. *Nucleic Acids Res.* **15**, 6273–6292 (1986).
29. Spangler, E. A., Ryan, T. & Blackburn, E. H. *Nucleic Acids Res.* **16**, 5569–5585 (1988).
30. Frankel, J., Nelsen, E. M., Jenkins, L. M. *Dev. Biol.* **58**, 255–275 (1977).
31. Frankel, J., Jenkins, L. M. & DeBault, L. E. *J. Cell Biol.* **71**, 242–260 (1976).
32. Runge, K. W. & Zakian, V. A. *Molec. cell Biol.* **9**, 1488–1497 (1989).
33. Gottschling, D. E. & Zakian, V. A. *Cell* **47**, 195–205 (1986).
34. Price, C. M. & Cech, T. R. *Genes Dev.* **1**, 783–793 (1987).
35. Price, C. M. & Cech, T. R. *Biochemistry* **28**, 769–774 (1989).
36. Raghuraman, M. K., Dunn, C. J., Hicke, B. J. & Cech, T. R. *Nucleic Acids Res.* **17**, 4235–4253 (1989).
37. Raghuraman, M. K. & Cech, T. R. *Cell* **59**, 719–728 (1989).
38. Blackburn, E. H. & Chiu, S.-S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2263–2267 (1981).
39. Budarf, M. L. & Blackburn, E. H. *J. biol. Chem.* **261**, 363–369 (1986).
40. Lundblad, V. & Blackburn, E. H. *Cell* (in the press).
41. Agard, D. A. & Sedat, J. W. *Nature* **302**, 676–681 (1983).
42. Shoeman, R. L., Wadley, S., Scherbarth, A. & Traub, P. *J. biol. Chem.* **263**, 18744–18749 (1988).
43. Shoeman, R. L., Pfundstein, S., Scherbarth, A. J. & Traub, P. *J. biol. Chem.* (in the press).
44. Kimmerly, W. J. & Rine, J. *Molec. cell Biol.* **7**, 4225–4237 (1987).
45. Diffley, J. F. X. & Stillman, B. *Nature* **342**, 24 (1989).
46. Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492 (1985).
47. Labhart, P. & Reeder, R. H. *Cell* **45**, 431–443 (1986).
48. Kingston, R. E. in *Current Protocols in Molecular Biology* (eds Ausubel, F. M. et al.) (Wiley, New York, 1988).

ACKNOWLEDGEMENTS. We thank J. W. Sedat for help with photomicrography, and our colleagues in the laboratory for comments on the manuscript. We are grateful to Y.-S. Jin for assistance and technical expertise in photography. This work was supported by the NIH (E.H.B.) and the NSF (J.D.B.).

LETTERS TO NATURE

X-ray reflection from cold matter in the nuclei of active galaxies

K. A. Pounds*, K. Nandra*, G. C. Stewart*,
I. M. George† & A. C. Fabian†

* X-ray Astronomy Group, Department of Physics and Astronomy,
University of Leicester, Leicester LE1 7RH, UK

† Institute of Astronomy, University of Cambridge,
Cambridge CB3 0HA, UK

THE evidence accumulated over the past few years for strong soft X-ray emission from active galactic nuclei^{1–3} has been interpreted as black body emission from the innermost stable region of an accretion disk feeding the putative black hole at the centre of the active nucleus, a view given strong support by the rapid variability of some soft X-ray components⁴. More recently, new X-ray data from the Exosat and Ginga satellites have revealed a second indicator of optically thick matter in the vicinity of the active nucleus, in the form of an iron K-fluorescence line at ≈ 6.4 keV (refs 5–8). We report the discovery of two further common features of continuum absorption and reflection, revealed in a composite spectrum from twelve Ginga observations of Seyfert-type active galactic nuclei. Most of these spectral features are shown to be well modelled by reprocessing of the hard X-ray power-law continuum in a slab (or perhaps a disk) of cold matter. There is also evidence for a substantial line-of-sight column of photoionized material.

The fact that active galaxies radiate at least 10% of their bolometric luminosity in the hard X-ray band (~ 2 –20 keV) was established over 10 years ago^{9–11}, and early spectral data showed this X-ray emission to be well described by a single power law, of energy index $\alpha \approx 0.7$ (refs 12, 13). Later measurements modified this view, with the discovery by Exosat that many Seyfert-type, emission-line active galactic nuclei (AGNs) had excess soft X-rays below ~ 1 keV (refs 1, 3) and with Einstein Observatory data showing similar results for the higher-luminosity quasi-stellar objects². The detection of rapid (of the order of hours) variability in the soft X-ray component, in several cases, supports its interpretation as thermal emission ($\sim 10^5$ K) from optically thick matter close to the central power source⁴. It is then natural to associate this matter with the primary accretion process feeding the active nucleus; it could exist in an accretion disk or in clouds, sheets or filaments, the chief requirement being for the material to be sufficiently dense to survive the strong heating close to the central nucleus¹⁴.

Further evidence for neutral matter in AGNs has recently been provided by the detection of an iron K-fluorescence line near 6.4 keV superimposed on the power-law continuum of several Seyfert-type AGNs^{5–8}. A second, more controversial, finding in Ginga and also in some Exosat data, is an apparent steepening of the power-law slope with increasing X-ray flux^{8,15}. Again, this is of particular interest in relation to theoretical predictions of slope changes arising from electron scattering

effects in a highly compact source¹⁶. An alternative explanation of the same data has been proposed, however, in the form of a partially ionized, or 'warm', absorber along the line of sight to the X-ray nucleus^{7,17}. Clarification of these more recent results has been hampered by the limited statistics of the available X-ray spectra, particularly above ~ 10 keV.

To attempt to redress this limitation, data from twelve Ginga observations of eight Seyfert-type AGNs have been superimposed, to produce a composite X-ray spectrum, Ginga-12. The objects chosen (Akn120, Mkn335, MCG-6-30-15, NGC3227, NGC4051, NGC5506, NGC5548 and NGC7314) are all bright in hard X-rays, individually having 'canonical' power-law slopes (from the same Ginga data) with an energy index ranging from $\alpha \approx 0.63$ to $\alpha \approx 0.85$ (with the exception of Mkn335 which had $\alpha \approx 1.1$). A simple power-law fit to the Ginga-12 data yielded a best-fitting energy index, $\alpha \approx 0.70$, in close agreement with the weighted mean slopes of the individual spectra, supporting the validity of the summing procedure. Of most interest are the data-minus-model residuals from this fit, shown in Fig. 1. In addition to the strong positive feature near 6 keV, there is a dip between ~ 8 and 12 keV and an excess above ~ 12 keV. Given that an iron K-fluorescence line is seen in each individual spectrum, together with some indications of K-edge absorption, it is natural to try a second spectral model, of a power law plus K-line and K-edge. The resultant fit has $\alpha \approx 0.64$ (with 90% confidence limits of 0.62–0.68), a small equivalent hydrogen absorbing column, $N_H \approx 4.1$ (3.0 – 5.3) $\times 10^{21}$ cm⁻², plus a K-line and edge as detailed below. The fit is remarkably good for such a composite spectrum, with a reduced chi-squared (χ^2_r) of 1.5 (for 38 degrees of freedom), compared with 9.0 (for 41 degrees of freedom) for the single power-law fit. The residuals of the best-fitting power law (Fig. 1) may now be quantified.

The Fe K-line has a mean (rest-frame) energy of 6.28 (6.07–6.49) keV. Although not well constrained, the suggestion of a mean energy less than that of neutral iron (6.4 keV) is interesting, given the expectation of a significant gravitational redshift in the fluorescence from matter close to the central hole¹⁸. The mean equivalent width of the line, of ~ 110 (75–150) eV, is also consistent with that expected from fluorescence from optically thick matter subtending a large solid angle (for example, an accretion disk) to the hard X-ray source¹⁹.

A large dip appears between ~ 8 and 12 keV. If this is due to Fe K-shell absorption, it corresponds to a large column density, of ~ 1.2 (0.7 – 1.9) $\times 10^{23}$ H atoms cm⁻² (assuming a solar abundance of iron), much greater than that implied by the small value of the Ginga-12 low-energy cutoff. One obvious explanation is that the line-of-sight material is strongly ionized, thereby having a reduced low-energy opacity. The best-fit edge energy, of ~ 7.9 (7.5–8.7) keV, supports this interpretation.

The third obvious feature in the residuals shown in Fig. 1 is an excess flux above ~ 12 keV, which can be modelled well between 10 and 35 keV by a flat power law, of energy index $\alpha \approx 0.45$. This is also seen in several of our individual spectra and provides the first clear evidence for a 'high-energy tail' in the continuum spectra of low-luminosity AGNs, consistent with

the 'reflection bump' recently predicted by several workers^{14,20}. If such a hard spectral feature is indeed a common property of AGNs, it offers an important new explanation for the spectral shape of the X-ray background radiation²¹.

All these features are highly suggestive of the modified spectrum expected by reprocessing of hard X-rays by reflection from cold material. To quantify this we have fitted a model spectrum, obtained from a Monte Carlo simulation of a power-law X-ray continuum irradiating an optically thick, semi-infinite slab of cold matter with solar elemental abundances. A line-of-sight 'warm' absorber was also included in the spectral fit, which was found to agree well with the Ginga-12 data, with $\chi_r^2 \approx 0.8$ (37

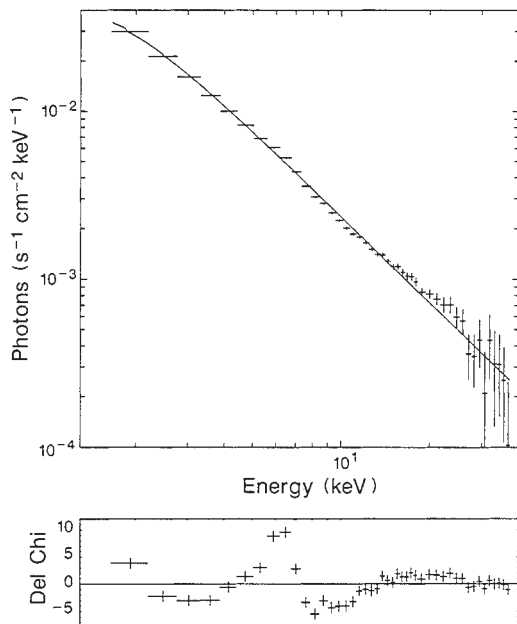


FIG. 1 Simple power-law fit to the Ginga-12 spectrum, together with a residual plot of the data minus model results.

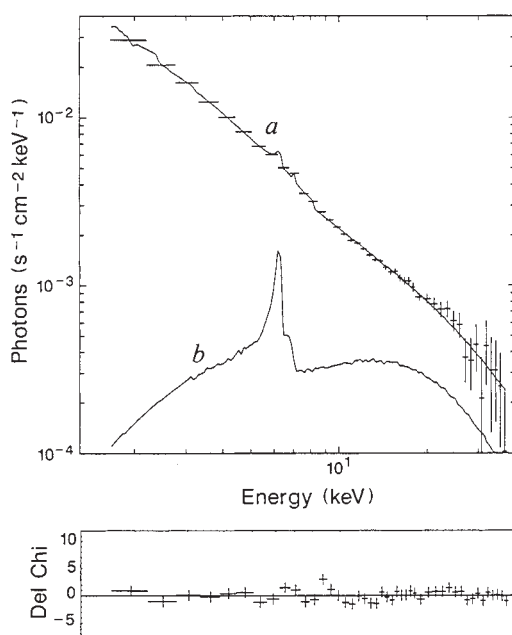


FIG. 2 *a*, Power law plus reflection and warm absorber model (as detailed in the text), together with residuals. *b*, Reflection component only.

degrees of freedom) (Fig. 2). The derived parameters were: $\alpha \approx 0.91$, $N_H \approx 6.0 \times 10^{21} \text{ cm}^{-2}$, a slab inclination (formally) of $\sim 9^\circ$ (but weakly constrained), with a residual warm Fe K-edge of $9 \times 10^{22} \text{ H atoms cm}^{-2}$ (solar abundance of iron) near $\sim 8.3 \text{ keV}$ and ionization parameter, $U (= Q/n_e r^2 c) \approx 5.1$, for the warm absorber (here Q is the number of ionizing photons, n_e is the electron density, r is the distance from the ionizing source and c is the velocity of light). Again, it is important to note that different subsets of the Ginga-12 sample gave essentially the same fit; furthermore, the overall model gave a satisfactory description of each individual data set. We conclude that the reflection and absorption features are common to the AGNs in our sample.

A fit was also tried without the warm absorber. Neglecting the lowest-energy channel, perhaps justified by the composite nature of the Ginga-12 spectrum, we found a good fit could again be obtained ($\chi_r^2 \approx 0.9$ for 39 degrees of freedom), with an enhanced ($1.5\times$) reflection (corresponding to a saucer-shaped disk), and an iron abundance twice that of the solar abundance. In effect, the enhanced iron line and reflection bump mimic the appearance of an absorption edge near $\sim 8 \text{ keV}$.

Both the reflected spectrum and the warm absorber can give rise to spectral variability. The reflected component is likely to smear out and delay rapid changes in the direct flux, such that the amplitude of rapid variability will be less for the iron line and above 10 keV , where the continuum reflection becomes significant. The warm absorber can change rapidly only if its density is high; a recombination timescale of $< 10^4 \text{ s}$, necessary in explaining the observed spectral changes, requires an electron density in the absorbing matter of $> 10^7 \text{ cm}^{-3}$. The ionization parameter inferred for this gas locates it within 10^{16} cm of the central hard photon source.

We draw attention to the attractive unification of several observed properties of AGNs offered by our interpretation of the new X-ray spectral features. The variable soft excesses and (possibly) redshifted Fe K-shell lines reveal cold, dense matter within 100 gravitational radii $r_g (=GM_H/c^2)$, of a central black hole of mass M_H ; the Fe K-line broadening found in some individual sources supports this view¹⁹. We now have quantitative evidence for the reflected continuum component from this cold matter. There is also evidence for highly photoionized gas along the line of sight and very close to the nucleus, with a Thomson depth of ~ 0.05 .

Finally, we note that the underlying power-law slope of the direct X-ray continuum in our reflection model is ~ 0.9 , steeper than the canonical value of 0.7 , suggesting that a reprocessed steeper continuum of the form discussed here is common, at least for Seyfert-type AGNs in the luminosity range of our sample, $L^{2-10} \approx 10^{42} - 10^{44} \text{ erg s}^{-1}$. This is important because most theoretical studies on the primary X-ray emission spectrum in AGNs have concentrated on mechanisms that yield $\alpha \approx 0.7$. $\square$

Received 16 October 1989; accepted 25 January 1990.

1. Arnaud, K. A. *et al. Mon. Not. R. astr. Soc.* **217**, 105-114 (1985).
2. Wilkes, B. J. & Elvis, M. *Astrophys. J.* **323**, 243-262 (1987).
3. Turner, T. J. & Pounds, K. A. *Mon. Not. R. astr. Soc.* **240**, 833-880 (1989).
4. Turner, T. J. & Pounds, K. A. *Mon. Not. R. astr. Soc.* **232**, 463-471 (1988).
5. Nandra, K. A., Pounds, K. A., Stewart, G. C., Fabian, A. C. & Rees, M. J. *Mon. Not. R. astr. Soc.* **236**, 39-46P (1989).
6. Pounds, K. A. *et al. Mon. Not. R. astr. Soc.* **240**, 769-784 (1989).
7. Nandra, K. A. *et al. Mon. Not. R. astr. Soc.* (in the press).
8. Matsuoka, M. *et al. Astrophys. J.* (submitted).
9. Pounds, K. A. *Ann. N.Y. Acad. Sci.* **302**, 361-385 (1977).
10. Elvis, M. *et al. Mon. Not. R. astr. Soc.* **183**, 129-157 (1978).
11. Tananbaum, H. *et al. Astrophys. J.* **223**, 74-81 (1978).
12. Rothschild, R. E. *et al. Astrophys. J.* **269**, 423-437 (1983).
13. Mushotzky, R. F. *Adv. Space Res.* **3**, 10-18 (1984).
14. Guilbert, P. W. & Rees, M. J. *Mon. Not. R. astr. Soc.* **233**, 475-484 (1988).
15. Perola, C. *et al. Astrophys. J.* **306**, 508-521 (1986).
16. Lightman, A. P. & Zdziarski, A. A. *Astrophys. J.* **319**, 643-661 (1987).
17. Yaqoob, T., Warwick, R. S. & Pounds, K. A. *Mon. Not. R. astr. Soc.* **236**, 153-170 (1989).
18. Fabian, A. C. *et al. Mon. Not. R. astr. Soc.* **238**, 729-736 (1989).
19. George, I. M., Nandra, K. A. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **242**, 28-31P (1990).
20. Lightman, A. P. & White, T. R. *Astrophys. J.* **335**, 57-66 (1988).
21. Fabian, A. C., George, I. M., Miyoshi, S. & Rees, M. J. *Mon. Not. R. astr. Soc.* **242**, 14-17P (1990).

Amorphization of cubic ice by ultraviolet irradiation

Akira Kouchi & Toshio Kuroda

Institute of Low Temperature Science, Hokkaido University, Sapporo 060, Japan

DESPITE current interest¹⁻⁸ in the effects of ultraviolet radiation and high-energy particles on interstellar ice and on solid mixtures of H₂O and other molecules, most attention has so far been paid to the initiation of chemical reactions, although it has been recognized that electron-beam bombardment will convert ice crystals to the amorphous form^{7,8}. Accordingly, we have studied the effect of ultraviolet radiation on ice, using electron diffraction to follow the structural changes that take place. We find that, below 70 K, ultraviolet irradiation transforms cubic ice (ice I_c) into amorphous ice (a-H₂O). This is a novel means of forming a-H₂O, supplementing previously known routes⁸⁻¹⁴. We infer that other high-energy rays or particles will amorphize ice crystals below 70 K. We conclude that crystalline ice in stellar atmospheres and on asteroids and planetary satellites will be rendered amorphous by ultraviolet radiation in times short compared with the timescale of stellar evolution.

A thin film of ice I_c, 1 µm in thickness, was prepared at 135 K by the condensation of H₂O vapour onto a metal substrate in a vacuum chamber at a pressure of 5×10^{-7} Pa. The substrate was cooled by a closed-cycle helium refrigerator, and was surrounded with a 10-K thermal insulator which was cooled by another refrigerator to prevent condensation onto the substrate of residual gas in the vacuum chamber. After deposition, the substrate was cooled to the desired temperature and the ice film was irradiated with ultraviolet radiation produced by a D₂ lamp with a MgF₂ window (Hamamatsu L879, wavelength 110–400 nm). The ultraviolet flux is $\sim 10^{12}$ photon cm⁻² s⁻¹, determined using a GaAsP Schottky diode with the window removed from the housing (Hamamatsu G1127-02). (Details of the diode are reported elsewhere¹⁵.) Structural changes in the ice film were examined *in situ* by reflection electron diffraction (electron energies of 20 keV).

We found that ice I_c was transformed below 70 K to a-H₂O by ultraviolet irradiation, whereas no change in structure was observed at temperatures above 70 K regardless of the irradiation time (up to 5 hours). Irradiation for 30 min to 1 h was

required for amorphization at temperatures between 50 and 70 K, whereas only several tens of seconds were required at 10 K. Figure 1 shows the reflection electron diffraction pattern of ice I_c (a) and ultraviolet-induced a-H₂O at 10 K (b). The diffraction pattern in Fig. 1b exhibits halo diffraction, which indicates that the irradiated ice is amorphous. On the basis of this diffraction technique, we can detect no structural differences between ultraviolet-induced a-H₂O and vapour-deposited a-H₂O at 10 K; moreover, the vapour pressure of the former is the same as that of the latter¹⁷. Our experimental method does not allow us to determine whether ultraviolet-induced a-H₂O corresponds to one of the forms of amorphous ice known previously^{12,13,18,19} or whether it is a new variant.

Figure 2 shows the temperature dependence of H₂ released from the ice films on ultraviolet irradiation. Two regimes are observed: the first is a sharp rise starting at ~ 17 K, resulting from desorption of H₂ from the ice surface^{20,21}, and the second, at temperatures higher than ~ 22 K, is attributed to H₂ release from the bulk. These results agree qualitatively with those of Laufer *et al.*²¹, who studied H₂ release from co-deposition of a 1:1 H₂-H₂O mixture. The release of H₂ increases in both regimes with increasing duration of irradiation, suggesting that the dissociation of H₂O molecules to H+OH and/or H₂+O may trigger the amorphization process. For equal duration of ultraviolet irradiation, the release of H₂ from vapour-deposited a-H₂O is greater than that from ultraviolet-induced a-H₂O. This is consistent with the idea that H₂ is formed more effectively in the micropores^{22,23} of the former than in the latter, in which initially there are no micropores.

The transformation temperature from a-H₂O to ice I_c on rewarming is 146 ± 1 K for both ultraviolet-induced and vapour-deposited a-H₂O, and does not depend on the duration of irradiation. If irradiation was continued during rewarming, however, the transformation temperature increased to 157 ± 1 K for both samples.

Lepault *et al.*⁷ found that crystalline ice (ice I_h and I_c) was transformed below 70 K into a-H₂O after irradiation with 100-keV electrons. They proposed that resistance of ice to amorphization above 70 K is due to a mechanism that restores the original structure after molecules have been displaced or dissociated by the electrons⁸. Below 70 K this mechanism is hindered so that the damage to the crystal structure is not repaired. Because the critical temperature for amorphization of ice I_c by ultraviolet irradiation is the same as that with electron irradiation, we infer that the restoration mechanism proposed by Dubochet and

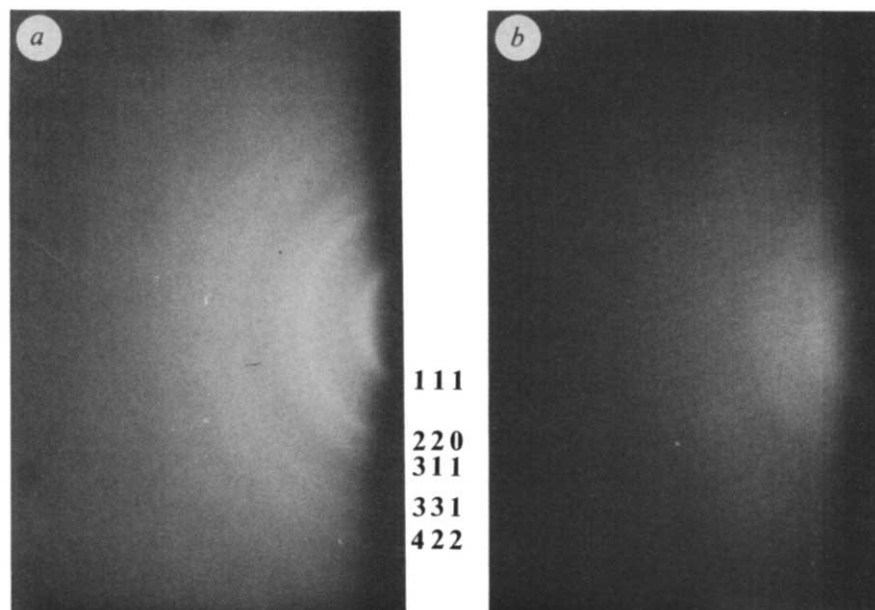


FIG. 1 Reflection electron diffraction patterns of a, ice I_c produced at 135 K, and b, ultraviolet-induced a-H₂O after 10 min of irradiation onto ice I_c at 10 K.

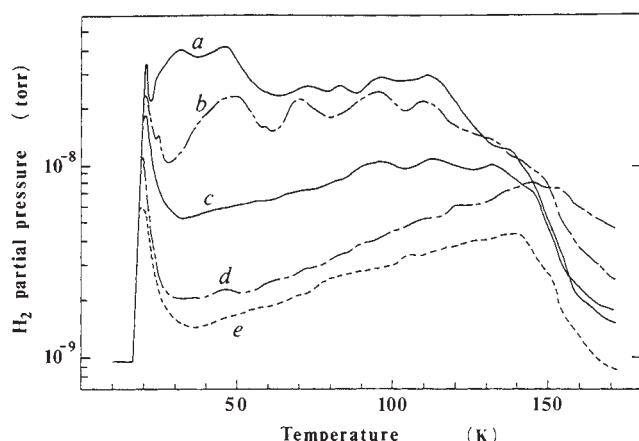


FIG. 2 Temperature dependence of H_2 release from ultraviolet-irradiated ice. *a*, *c*, vapour-deposited a- H_2O subjected to ultraviolet radiation; *b*, *d*, a- H_2O formed by ultraviolet irradiation onto ice I_c . The samples in *a* and *c* were $1 \mu m$ in thickness and were prepared by vapour deposition at a rate of 10^3 nm h^{-1} at 10 K. Durations of irradiation at 10 K were 10 min for *c*, *d* and 1 h for *a*, *b*. After irradiation, the substrate was warmed at a rate of 1 K min^{-1} . The release of H_2 from a- H_2O was observed with the partial-pressure mode of a quadrupole mass spectrometer (Edwards EQ80F) without calibration. Curve *e* is the release from non-irradiated a- H_2O deposited at 10 K, which must be due to H_2 impurity in the H_2O vapour and condensation of residual H_2 in the vacuum chamber. Differences between results *a*–*d* are clearly significant.

Lepault⁸ may also be operating in our experiments. By extension, we suggest that irradiation with other high-energy radiation or particles may also cause amorphization of crystalline ice below 70 K.

Our results have several implications for ice in extraterrestrial environments. The interstellar ultraviolet radiation fluxes in diffuse¹ and dense²⁴ molecular clouds are $\sim 10^8$ and $\sim 10^3 \text{ photon cm}^{-2} \text{ s}^{-1}$ respectively. The timescale equivalent to one hour of irradiation in the laboratory (that is, the time required for amorphization of ice I_c) is therefore only ~ 1 and $\sim 10^5 \text{ yr}$ for diffuse and dense clouds respectively. We may conclude that if crystalline ices are formed in interstellar clouds^{25,26}, they will be transformed by ultraviolet irradiation to a- H_2O following cooling below 70 K within a short time. In other words, the presence of a- H_2O in these environments does not necessarily imply that it condensed directly from the interstellar gas.

Mayer and Pletzer²² suggested that in the micropores of interstellar a- H_2O grains there is an increased probability of an encounter between two adsorbed H atoms, leading to the formation of H_2 . Our results (Fig. 2) suggest that ultraviolet radiation promotes the formation of H_2 in a- H_2O in interstellar clouds by supplying H atoms to the micropores from dissociation of H_2O molecules. We note that the size and number of micropores and thus the surface area of a- H_2O are changed with time by ultraviolet irradiation.

Because ice crystals (I_h and I_c) condensed at high temperature are energetically stable relative to a- H_2O , they persist below 70 K in the absence of radiation. Thus the presence of ice crystals in a stellar nebula that is thought to have a temperature $< 70 \text{ K}$ indicates an absence of ultraviolet radiation. Similar considerations apply to ice on the surfaces of asteroids and the satellites of the outer planets. □

Received 30 November 1989; accepted 30 January 1990.

- Hagen, W., Allamandola, L. J. & Greenberg, J. M. *Astrophys. Space Sci.* **65**, 215–240 (1979).
- Brown, W. L., Lanzetta, L. J. & Johnson, R. E. *Science* **218**, 525–531 (1982).
- Greenberg, J. M., van de Bult, C. E. P. M. & Allamandola, L. J. *J. phys. Chem.* **87**, 4243–4260 (1983).
- Bar-Nun, A., Herman, G., Rappaport, M. L. & Mekler, Yu. *Surf. Sci.* **150**, 143–156 (1985).
- Calcagno, L., Foti, G., Torrisi, L. & Strazzulla, G. *Icarus* **63**, 31–38 (1985).
- Rössler, K. *Radiat. Eff.* **99**, 21–70 (1986).
- Lepault, J., Freeman, R. & Dubochet, J. *J. Microsc.* **132**, RP3–4 (1983).
- Dubochet, J. & Lepault, J. *J. Phys. (Paris)* **45**, C7, 85–94 (1984).

- Soezats, M. G. & Rice, S. A. in *Water* Vol. 7 (ed. Franks, F.) 83–214 (Plenum, New York, 1982).
- Dubochet, J. & McDowell, A. W. *J. Microsc.* **124**, RP3–4 (1981).
- Mayer, E. & Brüggeller, P. *Nature* **298**, 715–718 (1982).
- Mishima, O., Calvert, L. D. & Whalley, E. *Nature* **310**, 393–395 (1984).
- Klug, D. D., Handa, Y. P., Tse, J. S. & Whalley, E. *J. chem. Phys.* **90**, 2390–2392 (1989).
- Hemley, R. J., Chen, L. C. & Mao, H. K. *Nature* **338**, 638–640 (1989).
- Krumrey, M. et al. *Appl. Opt.* **27**, 4336–4341 (1988).
- Kouchi, A. *Nature* **330**, 550–552 (1987).
- Kouchi, A. *J. Cryst. Growth* (in the press).
- Narten, A. H., Venkatesh, C. G. & Rice, S. A. *J. chem. Phys.* **64**, 1106–1121 (1976).
- Mishima, O., Calvert, L. D. & Whalley, E. *Nature* **314**, 76–78 (1985).
- Brackmann, R. T. & Fite, W. L. *J. chem. Phys.* **34**, 1572–1579 (1961).
- Lafer, D., Kochavi, E. & Bar-Nun, A. *Phys. Rev. B* **36**, 9219–9227 (1987).
- Mayer, E. & Pletzer, R. *Nature* **319**, 298–301 (1986).
- Schmitt, B., Ocampo, J. & Klinger, J. *J. Phys. (Paris)* **48**, C1, 519–525 (1987).
- Prasad, S. S. & Tarafdar, S. P. *Astrophys. J.* **267**, 603–609 (1983).
- Seki, J. & Hasegawa, H. *Astrophys. Space Sci.* **94**, 177–189 (1983).
- Kouchi, A. in *Dynamic Processes of Material Transport and Transformation in the Earth's Interior* (ed. Marumo, F.) (Terra, Tokyo, in the press).

ACKNOWLEDGEMENTS. We thank Professor H. Hasegawa for discussions. This work was partly supported by a grant-in-aid for scientific research on Priority Areas (Origin of the Solar System) from the Japanese Ministry of Education, Science and Culture.

Annual measurement of sea-ice thickness using an upward-looking sonar

Rick Hudson

Applied Microsystems Ltd, Sidney, British Columbia, Canada V8L 3S1

THE difficulties in collecting reliable data in polar regions are nowhere more acute than in sampling sea-ice growth and thickness. Yet ice dynamics and the forms the ice takes are of great importance to operators in polar conditions. Offshore rigs, harbours and icebreaking ships, for example, depend on reliable estimates of the sizes, distributions and mechanical properties of pressure ridges and undeformed floes. Here I report measurements of sea-ice thickness in the Canadian Beaufort Sea, using a bottom-mounted upward-looking sonar device operated for 250 consecutive days. Although carried out over a decade ago, the project is, as yet, the only successful experiment of its kind.

Under-sea profiles of ice ridges have been obtained from upward-looking sonars on submarines^{1,2}, and airborne downward-looking laser profiles have been combined³ with them in attempts to find a relationship between the underside of ridges (keels), where much of the mass is stored, and the ridge crests (sails), which are more easily measured. The results, unfortunately, do not give a time series from which annual or 10-year

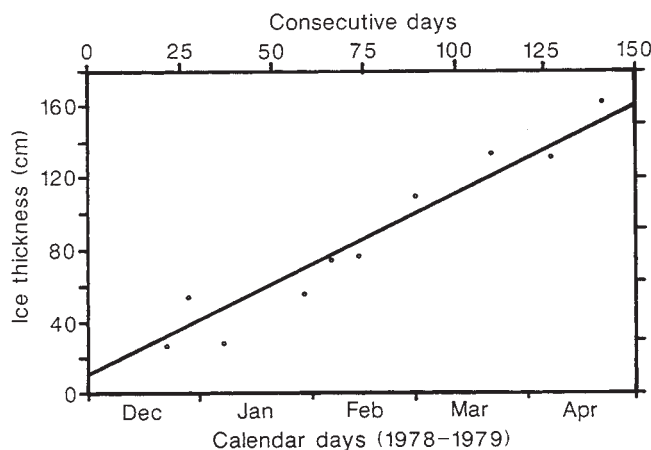


FIG. 1 Sea-ice thickness plotted against time. The data points are derived from ULS measurement of the undeformed ice sheets. The solid line represents a growth rate of 1 cm d^{-1} .

TABLE 1 Monthly keel depth results

Date	12 Dec- 12 Jan	13 Jan- 13 Feb	14 Feb- 17 Mar	18 Mar- 18 Apr	19 Apr- 20 May
Depth (m)	Number of events not exceeding this depth:				
0	0	0	0	0	0
2	0	0	0	0	0
4	42	28	12	31	88
6	160	36	100	135	43
8	78	12	50	98	180
10	50	1	17	50	75
12	21	0	40	23	61
14	7	0	76	11	40
16	3	0	19	7	15
18	0	0	6	1	2
20	0	0	20	1	4
22	0	0	6	0	3
Number of keels	361	77	346	357	511
Weighted keels	2,650	434	3,706	2,798	4,194
Mean weighted keel depth (m)	7.34	5.63	10.71	7.83	8.21
Mean undeformed ice sheet thickness (m)	0.40	0.72	1.02	1.30	1.53
Force unit width (kg s^{-2})	3,334	4,604	12,407	11,560	14,266
Mean driving stress (kPa)	81.8	62.7	119.3	87.2	91.5

maxima and return periods can be derived. Yet these are crucial to the design of offshore structures. And, for obvious reasons, submarines favour deep water regions, whereas most regions of active sub-sea exploration are along the continental margins.

A narrow-beam upward-looking sonar (ULS) and recorder were placed on the sea bed, in 30 m of water, at $70^{\circ}12.7' \text{ N}$, $133^{\circ}00.5' \text{ W}$ in the Canadian Beaufort Sea on 11 December 1978, and recovered 250 days later. The sounder sampled every 6 s with a range accuracy of 25 cm. Thus, as the sea ice drifted above, a continuous profile of the underside was recorded. These data, when recovered and analysed, have provided the first time series of sea-ice thickness in the transition ice zone, an area that lies between the shallow-water landfast ice to the south and the Arctic pack ice to the north.

The on-board logic of the recording unit determined keels using a 'fixed criterion' method: anything that exceeds a previous trough by more than 3 m was defined as a keel. This technique is independent of the water depth (which was unknown before the actual deployment), the thickness of the level ice, or the keel structure. After the data were recovered, a percentage exceedence or Rayleigh criterion was applied: if the draft of a keel exceeded the draft of neighbouring troughs by 50%, it was considered to be a keel. In general, a fixed criterion selects more of the large keels and the Rayleigh criterion selects smaller ones.

A stationary ULS is restricted to time-series measurements but spatial resolution can be derived by using the mean drift rate, measured by some other method. The direction of drift would be highly variable, of course, and areas of ice were probably rescanned occasionally. Over a period of 250 days, however, this does not affect the statistical base.

The growth rate of the undeformed ice was found to be lower than that measured by other means (such as drilling sample holes on a monthly basis). Two unexpected phenomena are evident in the plot of undeformed first-year floe thicknesses (in 25-cm steps) against time (Fig. 1). The ice growth from freeze-up in October until mid-December is erratic and minimal, despite normal temperatures from that time of the year. The thin ice, small pressure ridges and open water recorded by the ULS were confirmed by satellite imagery. The heat loss from ocean to atmosphere is two orders of magnitude larger for open leads than for thin ice⁴, and is usually associated with fog. This may partly explain the second anomaly: the growth rate was 1 cm d^{-1} , which is half that predicted by thermodynamic models^{5,6}. Other possible explanations for the low growth rate are a higher-than-average snow cover, which would have insulated the ice cover more effectively, or an above-normal heat flux in the sea, perhaps due to the Mackenzie River runoff before freeze-up.

In calculating the distribution of the keels I have made some adjustment for the effects of grounding close to the ULS late in the season (March) by removing such occurrences, which are obvious in the time series (Fig. 2). In Fig. 3 the accumulated time in days at which keels of a given depth were recorded (a valuable design parameter) is plotted against keel depth. When the data from grounded keels are removed, the trend is smooth.

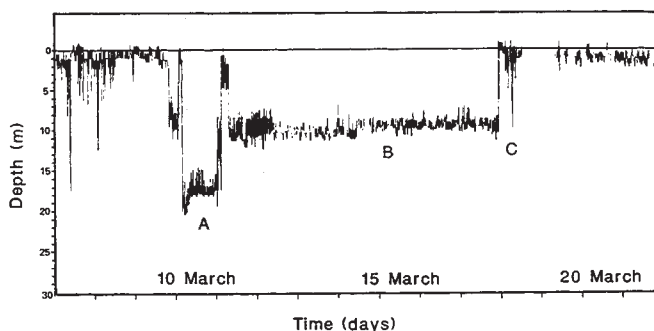
The number of ice keels (based on a Rayleigh criterion) of a given depth are shown in Fig. 4. Wadhams and Horne¹ have shown that for submarine transects the random draft distribution is of the form $n(D) dD = a \exp(-bD) dD$, where $n(D) dD$ is the number of keels of draft between D and $D + dD$, and a and b are experimentally derived constants. Using a keel-draft cutoff of 3 m and a total of 1,573 keels, the mean keel depth was found to be 5.4 m, and the best fit to the data (Fig. 4) gives the parameters $a = 2,089 \pm 700$ and $b = 0.408 \pm 0.025$. From this I calculate the return rates to be: maximum expected keel per year, 19 m; maximum expected keel per decade, 25 m; maximum expected keel per century, 31 m.

Grounding occurred close to the ULS at least twice during the period of study (see Fig. 2) and side-scan surveys of the area reveal a sea bed laced with recent ice scours. These data, based on a single year, may therefore not provide the complete picture and more conservative estimates of the return rates may be justified.

The ULS data may also be used to estimate pack-ice drift rates. From the results of a previous experiment⁷, the probability of finding a ridge 0–1 m high is 7.27 per km, which corresponds to 1.93 keels (of depth 0–1 m) per km assuming a sail-to-keel ratio of four. By comparison, the ULS data reduces to 2,089 0–1-m-high ridges per season. The ratio of these coefficients, 1,082, is a potential measure of the ice that has passed over the ULS during the winter. The 150 days of actual ice cover indicate ice velocities of 7 km d^{-1} , or 8 cm s^{-1} , which is in the range reported by other studies.

The driving force of the sea ice is important to designers of Arctic offshore platforms. It is, however, a difficult quantity to measure directly^{8,9}, and it is usually derived from observations

FIG. 2 A 15-day time series of ULS data. The horizontal axis is heavily compressed. In practice, the ULS sampled every 6 s. On the left, a typical cluster of pressure ridge keels are shown; in the centre, grounding of a deep keel nearby (water depth = 30 m) is visible (A, the keel is grounded; B, keel has moved but is still grounded; C, keel breaks free); on the right, an undeformed ice sheet is seen before ridges reappear.



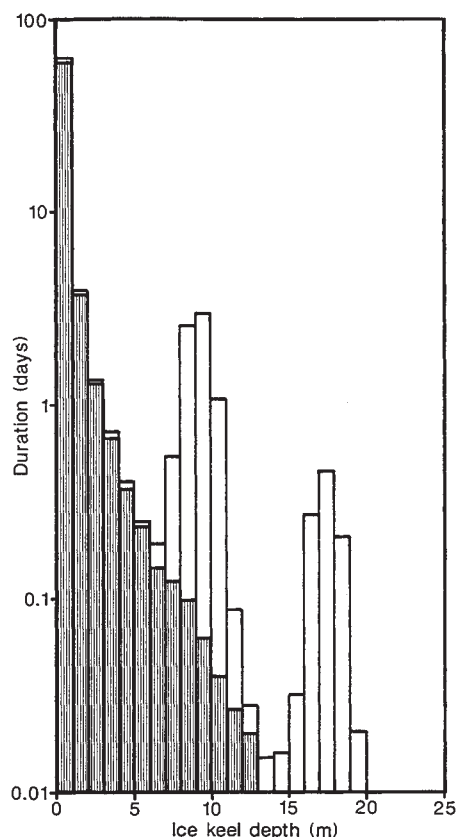


FIG. 3 Histogram showing the duration for which ice keels were at a particular depth. The anomalous 'peaks' at 9 m and 17 m are the result of local ice grounding. The shaded histogram shows the grounded data removed.

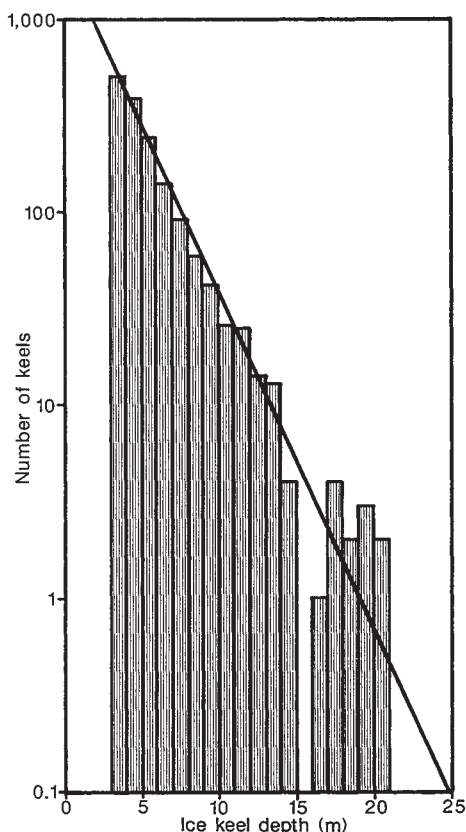


FIG. 4 Histogram of frequency of keel sizes for the winter season 1978/1979, using the Rayleigh keel identification criterion. The solid line is the best fit to a negative-exponential form.

of ice ridges seen in the field. Although the assumptions regarding failure mechanisms in the ice may be questioned, the large ULS data set provides a unique opportunity to estimate the driving forces involved.

In Table 1 the ULS data are divided into blocks of 32 consecutive days. From the accumulated total of keels observed during each period, I have calculated the mean weighted keel depth, and from the total time series, the mean undeformed ice thickness per period. Parmerter and Coon¹⁰ have shown that for a pressure ridge in isostatic equilibrium, the driving force F per unit width of ice sheet of density ρ_i , sail height H_s and sheet thickness t is

$$F = 0.5 \rho_i g H_s t$$

where g is the acceleration due to gravity.

Thus the mean driving stress (force per unit cross-sectional area), may be calculated using average measured values for t and derived values for H_s , and is 88.5 ± 10.2 kPa. F remains nominally constant throughout the winter season, suggesting a common deformation mode both for recently formed grey ice (20–40 cm thick) and seasoned first-year floes (1.2–1.7 m). □

Received 27 October 1989; accepted 9 January 1990.

1. Wadhams, P. & Horne, R. J. *J. Glaciol.* **25**, 401–424 (1980).
2. Wadhams, P. *Nature* **305**, 108–111 (1983).
3. Wadhams, P. *Phil. Trans. R. Soc. A* **302**, 45–85 (1981).
4. Badgley, F. I. *Proc. Symp. Arctic Heat Budget Atmos. Circ.* RM-5233-NSF, 267–268 (NSF, New Hampshire, 1966).
5. Maykut, G. A. & Untersteiner, N. *J. geophys. Res.* **75**, 1550–1552 (1971).
6. Thorndike, A. S. *J. geophys. Res.* **80**, 4501–4502 (1975).
7. Wadhams, P. *Arctic Ice Dynamics Joint Experiment Bull.* **33**, 1–52 (University of Washington, Seattle, 1975).
8. Pilkington, G. R. & Hudson, R. D. *Proc. Can. Congr. appl. Mech.* 42–47 (NRC, Moncton, 1981).
9. Danielewicz, B. W., Metge, M. & Hudson, R. D. *Proc. Port & Ocean Engng under Arctic Conditions Conf.* 629–642 (Arctic Petroleum Operators Association, Quebec, 1981).
10. Parmerter, R. R. & Coon, M. D. *J. geophys. Res.* **77**, 6565–6575 (1972).

ACKNOWLEDGEMENTS. I thank Dome Petroleum Ltd for support and the Arctic Petroleum Operators Association for permission to publish some of these data.

Aluminium as a depth-sensitive tracer of entrainment in submarine hydrothermal plumes

T. Lunel*, M. Rudnicki*, H. Elderfield* & D. Hydes†

* Department of Earth Sciences, Cambridge University, Cambridge CB3 0EZ, UK

† Institute of Oceanographic Sciences, Wormley, Godalming, Surrey GU8 5UB, UK

HYDROTHERMAL plumes formed by entrainment of ambient sea water into buoyant high-temperature vent fluids, are important features of ocean-ridge hydrothermal systems^{1,2}. An understanding of the mixing processes in plumes is essential for evaluating their role in modifying the chemical impact of hydrothermal activity on sea water and for calculating the heat and chemical budgets of ocean ridges. Central to all plume models is the rate-of-entrainment function which is used to determine the volume of water entrained by the buoyant plume at different depths^{3–5}. In the Pacific salinity can be used as a tracer of entrainment⁶ to a limited extent, but in the Atlantic Median Valley salinity does not vary appreciably near the sea bed. Here we show that aluminium can be used as a sensitive tracer as it varies markedly with depth in the immediate vicinity of hydrothermal vents and the aluminium signal from vent fluids is small, so that the aluminium entrainment signal is not masked.

Previous work at 45° N on the Mid-Atlantic Ridge has revealed high Al concentrations in Median Valley sea water⁷, but depth profiles do not indicate a direct correlation of Al with the

hydrothermal plume. This prompted measurement of Al concentrations during RRS *Discovery* cruise 176 to the Trans-Atlantic Geotraverse (TAG) area ($26^{\circ}\text{N } 44^{\circ}\text{W}$), August–September 1988. Samples were taken from Teflon-lined ‘Go-Flo’ bottles directly into acid-cleaned polyethylene bottles. Dissolved Al was determined using the Lumogallion fluorimetric method⁸. To avoid storage problems, analyses were carried out on board and begun within an hour of sample collection. A possible source of Al in the samples influenced by hydrothermal venting in Al-rich particles, so Al concentrations were measured in unfiltered samples and in samples of the same water filtered through acid-washed $0.4\text{-}\mu\text{m}$ and $0.2\text{-}\mu\text{m}$ Nuclepore filters. No significant differences were found, so the water samples were not filtered routinely, allowing more rapid analysis and avoiding possible contamination inherent in the filtration procedure.

Depth profiles of Al from within 200 m of the TAG hydrothermal mound (Fig. 1) show two features: (1) high values of Al about 300 m above the sea floor, associated with high particle concentrations as measured by nephelometry and indicative of the neutrally buoyant hydrothermal plume; and (2) background Al values, upon which the plume feature is imposed, that increase with depth. The background Al profile for the area was constructed by plotting the data from samples containing less than 2.5 nmol l^{-1} Mn from four stations within 1 km of the TAG mound, (Fig. 2); it shows a sharp increase of Al concentration over the bottom $\sim 500\text{ m}$ of the water column with a regional anomaly of $\sim 10\text{ nmol l}^{-1}$ relative to equivalent depths in the adjacent deep Atlantic ($20\text{--}30\text{ nmol l}^{-1}$)^{9,10}, suggestive of a bottom source.

Comparison of the observed Al values close to TAG (Fig. 1) with the concentrations predicted by the background Al profile shows that the size of the Al anomaly in the neutrally buoyant hydrothermal plume is between $10\text{--}14\text{ nmol l}^{-1}$. Measurements of Mn (ref. 1 and RRS *Discovery* cruise 176, unpublished data) in the neutrally buoyant plume show maximum concentrations less than 100 nmol l^{-1} which suggests that the end-member fluid is diluted by sea water by a factor of $10^4\text{--}10^5$. The highest Al values predicted by experimental determination of corundum solubility in pure water are $\sim 70\text{ }\mu\text{mol l}^{-1}$ Al (ref. 11). End-member hydrothermal fluids have not been successfully collected at TAG but hydrothermal Al values at MARK (Mid-Atlantic Ridge Kane) sites in the Atlantic and at sites on the East Pacific Rise are typically $5\text{--}20\text{ }\mu\text{mol l}^{-1}$ (refs 12, 13). These concentrations would produce an Al anomaly of only $0.1\text{--}1.0\text{ nmol l}^{-1}$ (or a maximum of $0.7\text{--}7\text{ nmol l}^{-1}$ if one makes the extreme assumption of an end-member concentration of $70\text{ }\mu\text{mol l}^{-1}$). Hence, there is an additional input of Al to the water column at plume height resulting from entrainment. Deep water, rich in dissolved Al, is entrained into the hydrothermal plume and taken to plume height producing an excess of Al above the concentrations defined by the background Al curve.

Previous workers^{6,14} have calculated the total amount of water entrained by hydrothermal plumes without ascribing the depth from which the water is entrained. The size of the Al signal produced by entrainment of deep water is determined by the rate-of-entrainment function (see Fig. 3 legend), and the background Al curve. Thus, by comparing the size of the Al anomaly predicted by the model with the measured Al anomaly, one can

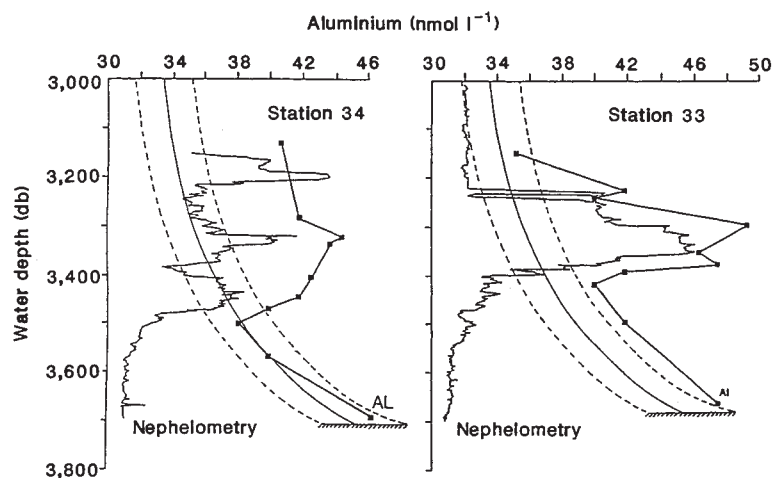


FIG. 1 Aluminium depth profiles (filled squares) 200 m from the TAG hydrothermal mound, showing correlation with the continuous record of particle concentration measured by nephelometry (in relative units). Units of water depth are decibars; $1\text{ db} \approx 1\text{ m}$ depth.

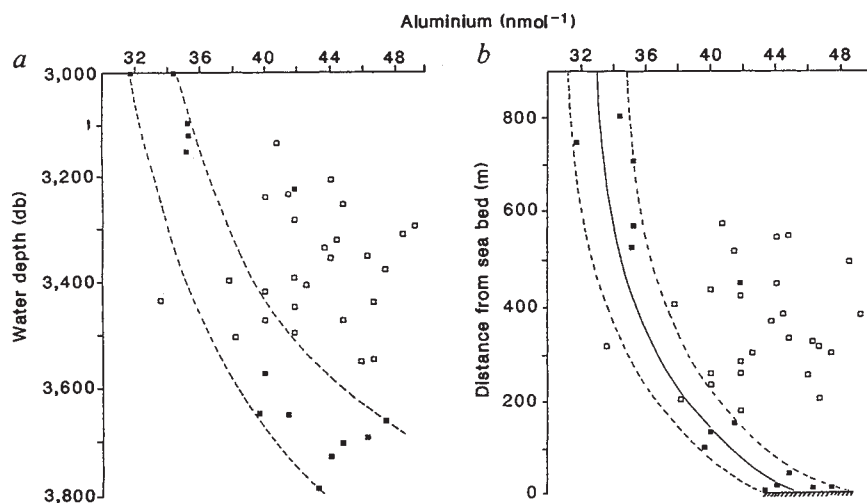


FIG. 2 A composite of four depth profiles of dissolved Al taken within radius of 1 km of the TAG hydrothermal mound. Filled squares denote samples for which measured Mn concentrations (RRS *Discovery* cruise 176, unpublished data) are less than 2.5 nmol l^{-1} and these are used to establish a background Al curve for the TAG area, as they have a small contribution from hydrothermal end-member fluid. (The dashed lines define the envelope for background Al concentrations, and the solid line represents the background Al curve which is used in the entrainment model.) The background curve is better constrained when the data are plotted, as shown, as a function of distance from the sea bed, rather than water depth.

FIG. 3 Application of the plume model to Al. Hatched areas indicate the total amount of Al that can be entrained into the rising hydrothermal plume. The rate of entrainment function is superimposed onto this to indicate the relative amounts of Al entrained at each depth. This function is calculated by setting up initial values for volume, momentum and density deficiency using estimated TAG values^{2,6,14} for radius of vent (5–15 cm), exit velocity of the fluid (0.4–0.8 m s⁻¹), and ambient density gradient (5 × 10⁻⁵ kg m⁻⁴). Changing these variables changes the height that the plume attains in the water column but does not change the shape of the rate of entrainment function, and therefore does not change the Al budget. To allow for the nonlinearity of the seawater temperature-density relation, the initial volume of hydrothermal fluid is diluted by a factor of nine with ambient sea water¹⁵. The fourth-order Runge-Kutta finite-difference method is used to generate values of radius, velocity and density deficiency over the entire plume height by conserving the volume, momentum and density deficiency of the system as a whole, while increasing the plume height by small increments. The rate of entrainment function is then used to calculate the amount of Al entrained at each discrete depth

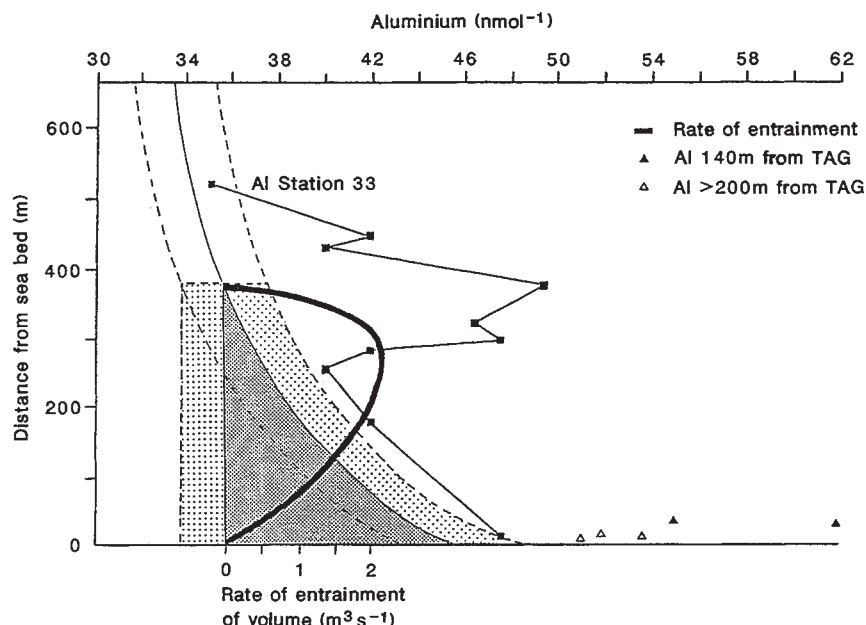
$$\text{Aluminium concentration at plume height} = \frac{(\sum_0^i (F_i [Al]_i)) + F_n [Al]_n}{(\sum_0^i (F_i)) + F_n} = 42 \text{ nmol l}^{-1} \text{ Al}$$

test the validity of the entrainment function. The sensitivity of this test is a result of the steep Al concentration gradient in this area.

The Al anomaly was determined using a plume model for a linearly stratified environment⁴ (Fig. 3). A modified expression for the initial density difference between the plume and ambient sea water was used¹⁵ which accounts for the nonlinear variation of seawater density with temperature by instantaneously diluting the emerging hydrothermal plume by a factor of nine by entraining ambient sea water. Use of this modified expression also slightly increases the size of the Al anomaly at plume height as the modification to the model incorporates more Al-rich bottom water. When this modified entrainment function is used to weight the contributions from the background Al curve at different heights, the concentration generated at plume height is 42 ± 2 nmol l⁻¹ Al (including an assumed contribution of 1 nmol l⁻¹ from the diluted hydrothermal end-member), an anomaly of ~7 nmol l⁻¹. It therefore seems that this axisymmetric plume model can account for ~80–90% of the total Al signal and 50–70% of the Al anomaly.

To test the entrainment model further, the model was applied to the Si data for the same sites. The entrainment model accounted for 36 ± 0.5 μmol l⁻¹ Si at plume height, 90–100% of the total Si signal at these sites. The contribution of Si from diluted end-member hydrothermal fluids¹³ is a larger proportion of the total Si anomaly than for Al, ~2 μmol l⁻¹ of a total anomaly of 3–4 μmol l⁻¹, and the background Si concentration does not increase with depth to the same extent as Al, so it is a less sensitive tracer of entrainment.

The fact that we cannot account for all of the Al or Si anomaly may, in part, reflect deficiencies in the plume model or the data used and must be significant because concentrations in the background profiles do not reach the values found at plume height. The large plume signal could be because the background profiles immediately adjacent to the vent site are higher in concentration than that used. Indeed, five samples from the bottom 50 m of the water column close to the TAG mound, showed Al values between 50–61 nmol l⁻¹, the two highest values recorded being 140 m from TAG (Fig. 3). But when these higher bottom-water concentrations are used in the model, the Al



where F_i is the flux of water entrained into the plume at height i ; Al_i is the Al concentration at height i ; F_n is the hydrothermal end-member fluid from the vent; and Al_n is the Al concentration in the hydrothermal end-member.

concentration at plume height is increased only slightly, to 43 ± 2 nmol l⁻¹. For the present plume model to account for all the Al signal, the elevated values must extend well up into the water column to the area around 200–250 m from the sea floor, where most of the entrainment takes place. This may occur by means of gravity currents, producing a sheath of Al-rich water flowing down the plume margins³.

The buoyant hydrothermal plume also contains resuspended sediment particles which are carried upwards by the flux of water^{15,16}. The release of dissolved Al by resuspension of sediment has been proposed for other areas^{17,18} and a similar process might be taking place here to augment the elevated Al concentrations at plume height.

The small discrepancies between the Al values predicted by the model and those observed here may provide further refinements in the future. However, this study shows that dissolved Al concentrations can be used to investigate the entrainment function currently used in plume models, providing an advance on earlier work using salinity⁴, and shows that the models are consistent with the results from a depth-sensitive tracer of entrainment such as Al. □

Received 23 October 1989; accepted 23 January 1990.

1. Klunkhammer, G. P., Elderfield, H., Greaves, M. R., Rona, P. A. & Nelsen, T. A. *Earth planet. Sci. Lett.* **80**, 230–240 (1986).
2. Speer, K. G. & Rona, P. A. *J. geophys. Res.* **94**, 6213–6220 (1989).
3. Carey, S. N., Sigurdsson, H. & Sparks, R. S. J. *J. geophys. Res.* **93**, 15,314–15,328 (1988).
4. Morton, B. R., Taylor, G. & Turner, J. S. *Proc. R. Soc. A* **234**, 1–23 (1956).
5. Turner, J. S. *J. Fluid Mech.* **173**, 431–471 (1986).
6. Lupton, J. E., Delaney, J. R., Johnson, H. P. & Tivey, M. K. *Nature* **316**, 621–623 (1985).
7. Hydes, D. J., Statham, P. J. & Burton, J. D. *Sci. Total Envir.* **49**, 133–145 (1986).
8. Hydes, D. J. & Liss, P. S. *Analyst* **101**, 922–931 (1976).
9. Hydes, D. J. *Science* **205**, 1260–1262 (1979).
10. Measures, C. I., Edmond, J. M. & Jickells, T. D. *Geochim. cosmochim. Acta* **50**, 1423–1429 (1986).
11. Ragnarsdottir, K. V. & Walther, J. V. *Geochim. cosmochim. Acta* **49**, 2109–2115 (1985).
12. Bowers, T. S. *et al. J. geophys. Res.* **93**, 4522–4536 (1988).
13. Campbell, A. C. *et al. Nature* **335**, 514–518 (1988).
14. Little, S. A., Stolzenbach, K. D. & von Herzen, R. P. *J. geophys. Res.* **92**, 2587–2596 (1987).
15. Turner, J. S. & Campbell, I. H. *Earth planet. Sci. Lett.* **86**, 85–92 (1987).
16. Converse, D. R., Holland, H. D. & Edmond, J. M. *Earth planet. Sci. Lett.* **69**, 159–175 (1984).
17. Lunel, T. thesis, Univ. Cambridge (1989).
18. Mackin, J. E. & Aller, R. C. *Geochim. cosmochim. Acta* **48**, 281–297 (1984).

ACKNOWLEDGEMENTS. We thank the Captain, crew and scientists of RRS *Discovery* cruise 176 for their assistance, in particular B. Dickie, L. Godfrey, M. Greaves and A. Mitra. This work was supported by NERC.

Right-lateral shear and rotation as the explanation for strike-slip faulting in eastern Tibet

Philip England* & Peter Molnar†

* Department of Earth Sciences, Parks Road, Oxford, OX1 3PR, UK

† Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

THE convergence of two continents can be accommodated by crustal thickening and by lateral transport of crust out of the path of the converging continents^{1–3}. The principal features suggesting that lateral transport is important are major strike-slip faults striking roughly orthogonally to the orientation of convergence between the continents. Displacement on such faults is imagined to allow lateral transport of material with respect to the converging continents, possibly accounting for a large fraction of the convergence between the two continents. The most spectacular example of such strike-slip faulting is in the eastern part of the Tibetan Plateau, where three major left-lateral faults, with slip rates of $>10 \text{ mm yr}^{-1}$, and several smaller roughly parallel faults dominate the regional strain field^{2,3}. Cobbold and Davy⁴ suggested that these faults may be rotating in a clockwise fashion. Here we put bounds on that rate of rotation, and conclude that the image of lateral transport on such faults², known also as ‘continental escape’, ‘extrusion’, or ‘expulsion’, is an illusion, and that instead the left-lateral slip on east-striking planes in eastern Tibet is a manifestation of north-striking right-lateral simple shear. If this conclusion is correct, the east-striking left-lateral faults and the crustal blocks between them are rotating clockwise at $1\text{--}2^\circ \text{ Myr}^{-1}$, the east–west dimension of eastern Tibet is shortening at $10\text{--}20 \text{ mm yr}^{-1}$, and little material is moving eastward out of India’s path into Eurasia by left-lateral simple shear.

Three major faults dominate the tectonics of eastern Tibet (Fig. 1). Offsets of late Quaternary moraines and erosional features imply Holocene left-lateral slip of at least 20 mm yr^{-1} along the Altyn Tagh fault in western Tibet; the slip rate is probably close to 10 mm yr^{-1} near its eastern extremity⁵. The 30-km offset of ground moraine from its source along the Kunlun, or Xidatan-Tuosuohu-Maqu, fault implies that the Quaternary rate of left-lateral slip along this fault exceeds 10 mm yr^{-1} (ref. 6). Similar offset features⁷ and displacements associated with earthquakes in this century⁸ suggest a Holocene rate of at least 10 mm yr^{-1} for the Xianshuhe fault (Fig. 1). Holocene and Quaternary slip rates on at least two other, more minor, roughly east-striking left-lateral faults in eastern Tibet are $5\text{--}6 \text{ mm yr}^{-1}$ (ref. 9) and $6\text{--}10 \text{ mm yr}^{-1}$ (refs 10, 11). Finally, left-lateral displacements of one to a few metres can be associated with two major earthquakes ($M > 7$) on another two, roughly east–west trending faults in eastern Tibet in this century⁸. Thus, the general impression (see Fig. 2a), that the area between 95° E to 105° E and 30° N to 40° N is undergoing major left-lateral shear on roughly east–west-trending planes, is virtually inescapable.

The logical step from recognizing such left-lateral shear to inferring that material is systematically displaced eastward by slip on these faults^{2,3,8,12–14} relies on an implicit, and generally unrecognized, assumption that the faults separating blocks of crust do not rotate. Rotations of crustal blocks do occur by slip on curved strike-slip faults in eastern Tibet^{14,15}, but observations of faulting alone cannot quantify, or even detect, rotations of the faults themselves. If the strike-slip faults rotate about vertical axes, with respect to an external reference frame such as Eurasia, India or southeast China, then the crustal blocks that they separate must also rotate by this additional amount, which can be large⁴. Such rotations, undetected by structural geology, have

been demonstrated clearly by palaeomagnetic results from several other regions of continental deformation^{16–18}.

Suppose that eastern Tibet underwent shear in a right-lateral sense along a north–south-trending zone. In such a zone, the roughly east-trending blocks and the faults separating them (Fig. 1) would rotate clockwise, slip on the faults would be left-lateral, and the eastern and western margins of the zone would approach each other¹⁹ (Fig. 2b).

Common sense, unburdened by a knowledge of the faulting in eastern Tibet, would lead to the inference of such north-trending, right-lateral simple shear. The eastern part of India moves nearly due north at $50\text{--}60 \text{ mm yr}^{-1}$ with respect to Eurasia²⁰. India underthrusts southern Tibet, also nearly due north^{15,21}, at between 10 and 25 mm yr^{-1} (refs 22, 23). Hence, southern Tibet, just north of eastern India, moves northward at $25\text{--}50 \text{ mm yr}^{-1}$ with respect to Eurasia. Southeast China, the area east of eastern Tibet, is surely not rigidly attached to Eurasia, but the following argument shows that its northward component of velocity with respect to Eurasia is a small fraction of that of southern Tibet ($\leq 10 \text{ mm yr}^{-1}$). All known major active structures between southeast China and eastern Siberia reflect strike-slip or normal faulting^{2,3,24,25}. Slip on these structures may allow southeast China to move east or southeast with respect to Eurasia^{8,12,14,26} but would not permit a northward translation of southeast China with respect to Eurasia. The western part of southeast China could be moving northwards with respect to Eurasia, whereas the northern part would not if there were rapid ($>1^\circ \text{ Myr}^{-1}$) clockwise rotation of southern China about an axis within southeast China. The hypothesis of a rapid clockwise rotation of southern China is difficult to test, not least because the extent of the block containing southern China is hard to ascertain. If that block extended as far south as the Banda arc then the hypothesis would seem to be incorrect, because the slip vectors of earthquakes in the western Banda arc²⁷ would imply anticlockwise rotation of the block. In the absence of a more definitive test, we presume that rapid clockwise rotation of southeast China does not occur.

With that assumption, the northward component of southern Tibet’s velocity with respect to western southeast China should be at least 15 mm yr^{-1} (and could be substantially $> 30 \text{ mm yr}^{-1}$ and eastern Tibet is undergoing a significant component of regional north-trending, right-lateral simple shear. Note that such right-lateral shear is found in both numerical and laboratory experiments of deformation of a variety of media subjected to an indenting boundary condition, such as India applies to Asia^{4,12,14,26,28–34}.

We may test the hypothesis that north–south dextral shear is distributed across the region by using the difference across the zone in the north–south component of velocity to calculate the rate of rotation of the blocks and the rate and sense of shear on their bounding faults. Consider a rectangular region encompassing eastern Tibet, roughly $1,200 \text{ km}$ in its north–south dimension and $1,000 \text{ km}$ east–west (Fig. 1). Let us ignore the likely variations in the style of deformation and in the orientations of the faults, treating the deformation as homogeneous throughout the region. (We will discuss the complexities of the deformation elsewhere.) The rate of rotation of the blocks and the faults bounding them is given by the north–south component of the velocity difference across the zone, divided by the east–west width of the zone ($\sim 1,000 \text{ km}$). Above, we estimate that the north–south component of the velocity difference is $15\text{--}30 \text{ mm yr}^{-1}$. The rate of rotation due to this velocity difference is between 1 and 2° Myr^{-1} . The corresponding cumulative rate of left-lateral slip on east–southeast to southeast-trending faults is $20\text{--}50 \text{ mm yr}^{-1}$ over the whole zone, or roughly 10 mm yr^{-1} on each of the major faults. Hence, the calculated cumulative left-lateral shear agrees with the sum of measured slip rates on faults in eastern Tibet.

Because the major faults in this region are all vertical strike-slip faults, shortening of one dimension must be associated with

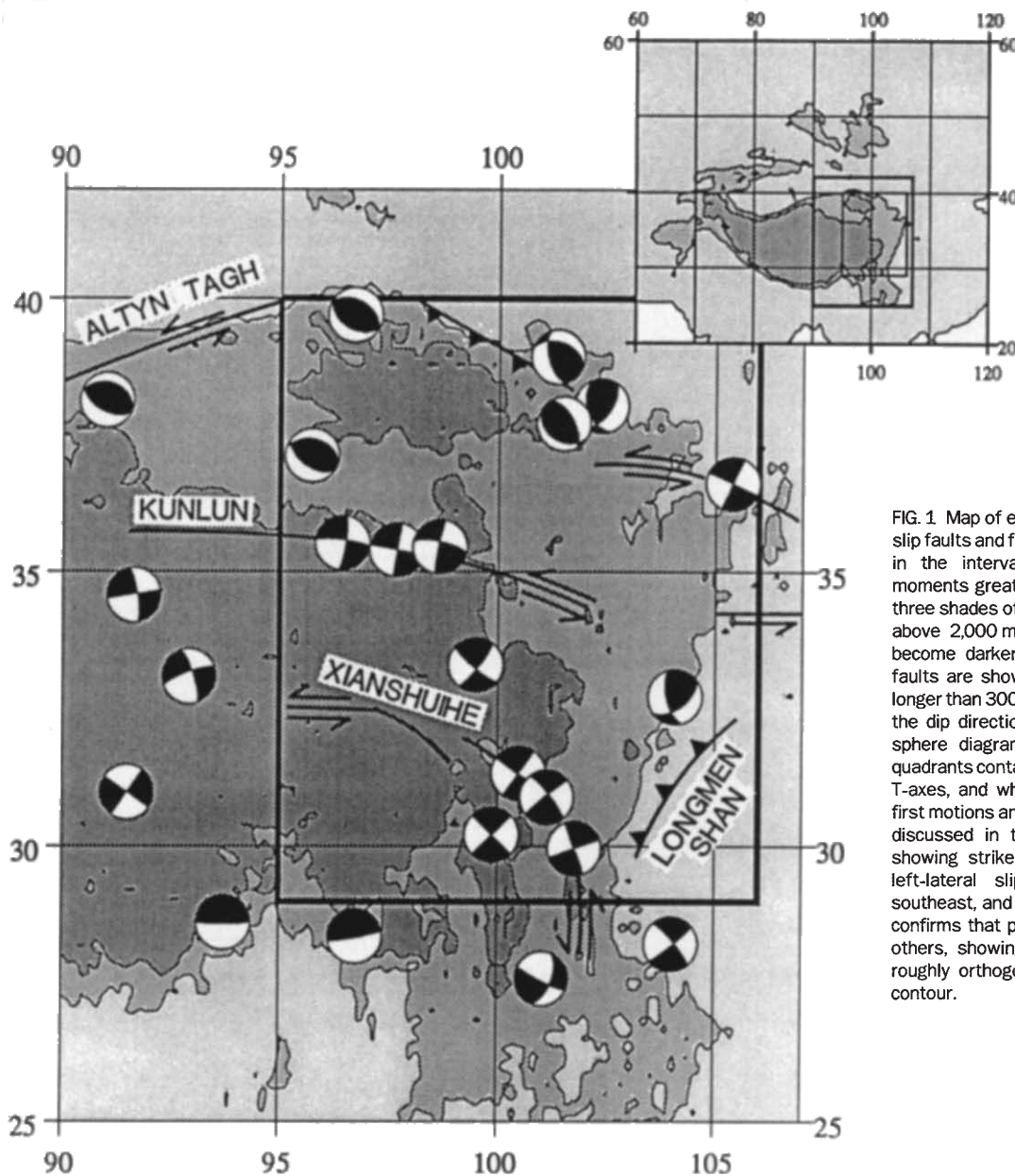


FIG. 1 Map of eastern Tibet, showing major strike-slip faults and fault-plane solutions of earthquakes in the interval 1900 to 1987 with seismic moments greater than 10^{18} N m (refs 8, 15). The three shades of grey indicate land above sea level, above 2,000 m and above 4,000 m; the shades become darker with increasing elevation. Major faults are shown by dark black lines; only faults longer than 300 km are included, and teeth indicate the dip direction of thrust faults. In lower-hemisphere diagrams of fault-plane solutions, dark quadrants contain compressional first motions and T-axes, and white quadrants contain dilatational first motions and P-axes. All solutions in the region discussed in this paper (95° – 105° E, 29° – 40° N) showing strike-slip faulting are consistent with left-lateral slip on a plane trending east-southeast, and for most of them, surface faulting confirms that plane to be the fault plane. For the others, showing thrust faulting, the P-axes are roughly orthogonal to the trend of the 2,000-m contour.

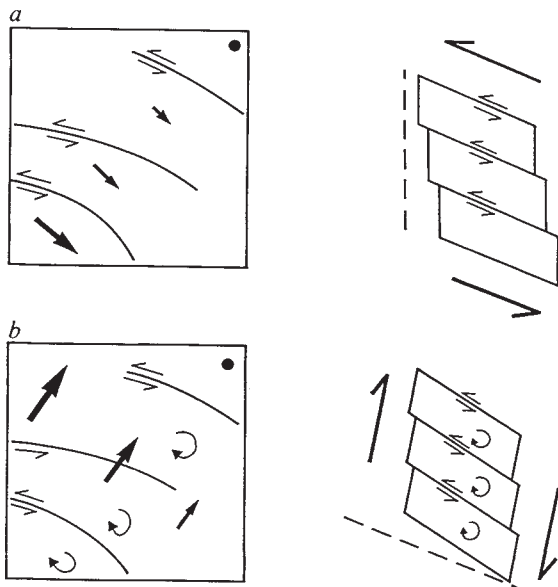


FIG. 2 Sketches of contrasting interpretations of the faulting in eastern Tibet. For *a* and *b*, the left figure shows locations and senses of slip on idealized faults with the inferred velocity field corresponding to the simple interpretation on the right. The velocities are drawn with respect to a fixed point in the upper right-hand corner. In *a*, as slip occurs, the faults do not rotate with respect to the surrounding material. Left-lateral slip leads to an east-southeastward translation of material on the south side of each fault relative to material on the north side. The sum of slip rates on the faults would give the rate at which material in the south is translated east-southeastward with respect to the material in the north. In *b*, the eastern boundary of the region defines the fixed reference frame, and the left-lateral slip is a manifestation of north-trending right-lateral shear of the region and clockwise rotation of the blocks between the left-lateral faults. Note that for the east-southeasterly strikes of the faults, the continued rotation and slip on these faults leads to a shortening of the east-west dimension and a lengthening of the north-south dimension^{4,19,35,36}.

perpendicular extension^{19,35,36}. Clearly, this shortening is oriented neither perpendicular to nor parallel to the strike-slip faults. Most of the left-lateral faults, south of the Altyn Tagh fault, strike east-southeast or southeast (110°–135°), and consequently, the clockwise rotation of blocks leads to a shortening of the east–west dimension and a lengthening in the north–south dimension⁴. The rate of east–west shortening is given by the north–south component of the velocity difference divided by the tangent of the strike of the faults. Hence the western margin of eastern Tibet (near 95° E) should move eastwards with respect to the eastern margin (near 105° E) at about 10–20 mm yr⁻¹.

We may now investigate quantitatively the question of whether southern China is², or is not⁴, being displaced eastward with respect to Eurasia by sinistral slip on the major faults in Eastern Tibet. The western margin of eastern Tibet is moving eastwards with respect to Eurasia at an estimated rate of between 10 and 30 mm yr⁻¹ (see for instance, refs 13, 15 and 26). If the velocity field of Fig. 2a is accepted we must add 20–40 mm yr⁻¹ to this rate to calculate the rate of eastward displacement of southeasternmost Tibet (the Longmen Shan in Fig. 2) with respect to Eurasia. By contrast, if the velocity field of Fig. 2b is accepted, we must subtract 10–20 mm yr⁻¹ to calculate that rate. If the first calculation (Fig. 2a) is performed, it seems that easternmost Tibet (and presumably southern China) is moving eastward with respect to Eurasia at several centimetres a year. If the second calculation (Fig. 2b) is performed, one concludes that southern China may be moving east with respect to Eurasia by at most a couple of centimetres a year—or may, in fact, not be moving eastward with respect to Eurasia at all. Thus the image² of rapid eastward ‘expulsion’ or ‘extrusion’ of southern China with respect to Eurasia cannot be sustained if the faults in eastern Tibet are rotating clockwise at more than about 1° Myr⁻¹—as our calculation of the rate of north–south dextral shear of the region suggests they are.

One might ask why the surface crust in a zone of north-trending, right-lateral shear deforms by displacements and rotations of blocks of crust bounded by roughly east-striking faults. Pre-existing weaknesses that formed both before India collided with Eurasia and later, when its penetration caused north–south shortening of southern Eurasia, probably trended roughly easterly. The palaeogeography of southern Eurasia³⁸ reveals several east-trending Mesozoic sutures across Tibet, and both numerical^{28–34,38} and laboratory⁴ experiments of continuous media in a gravity field show shortening in a north-to northeast-trending zone. The change from largely crustal shortening in Tibet to east–west extension in late Cenozoic time should have enhanced the eastward component of displacement³⁰ and might have encouraged left-lateral slip on these pre-existing weak zones. This explanation suggests that left-lateral faulting in eastern Tibet became significant only after east–west extension of Tibet began. If this is so, the rotations suggested here would be on the limit of detection by palaeomagnetic means, but will eventually be detectable by space geodetic techniques. Should the rotations and right-lateral shear in eastern Tibet prove to be difficult to measure, either with palaeomagnetism or with space geodesy, then an alternative is to test the basic assumption that the western margin of southeastern China does not move northward with respect to Eurasia at more than ~10 mm yr⁻¹.

The absence, hitherto, of inferences of clockwise rotation of faults and of north-trending right-lateral shear in eastern Tibet calls attention to a general aspect of continental tectonics. As stated clearly by McKenzie and Jackson^{19–23,35–37}, structural geology and seismology share a fundamental weakness in the study of tectonics, because they are poorly equipped to detect rotations of faults about vertical axes. It would be folly to suggest that the orientations and rates of slip on major faults are not important in describing the kinematics of regional deformation, but a description of deformation derived solely from faulting, however how high the rates of slip, can give a very misleading

impression of the velocity field of deforming continental lithosphere. □

Received 17 October 1989; accepted 22 January 1990.

- McKenzie, D. *Geophys. J. R. astr. Soc.* **30**, 109–185 (1972).
- Molnar, P. & Tapponnier, P. *Science* **189**, 419–426 (1975).
- Tapponnier, P. & Molnar, P. *J. geophys. Res.* **82**, 2905–2930 (1977).
- Cobbald, P. R. & Davy, P. *Bull. Geol. Inst. Uppsala* **14**, 143–162 (1988).
- Peltzer, G., Tapponnier, P. & Armijo, R. *Science* **246**, 1285–1289 (1989).
- Kidd, W. S. F. & Molnar, P. *Phil. Trans. R. Soc. Lond. A* **327**, 337–363 (1988).
- Allen, C. R. et al. *Geol. Soc. Am. Bull.* (in the press).
- Molnar, P. & Deng Qidong *J. geophys. Res.* **89**, 6203–6227, (1984).
- Peltzer, G. et al. *J. geophys. Res.* **93**, 7793–7812 (1988).
- Burchfiel, B. C. et al. *Tectonics* (in the press).
- Zhang Peizhen. *Quat. Res.* **30**, 151–164, (1988).
- Tapponnier, P., Peltzer, G., Le Dain, A. Y., Armijo, R. & Cobbald, P. *Geology* **10**, 611–616 (1982).
- Armijo, R., Tapponnier, P. & Han Tonglin. *J. geophys. Res.* **94**, 2787–2838, (1989).
- Peltzer, G. & Tapponnier, P. *J. geophys. Res.* **93**, 15,085–15,117, (1988).
- Molnar, P. & Lyon-Caen, H. *Geophys. J. Int.* **99**, 123–153, 1989.
- Beck, M. E. *Am. J. Sci.* **276**, 694–712, (1976).
- Luyendyk, B. P., Kameling, M. J., Terres, R. R. & Hornafius, J. S. *J. geophys. Res.* **90**, 12,454–12,466 (1985).
- Kissel, C. & Laj, C. (eds) *Paleomagnetic Rotations and Continental Tectonics* (Kluwer, Dordrecht, 1989).
- McKenzie, D. & Jackson, J. *Earth planet. Sci. Lett.* **65**, 182–202, (1983).
- Minister, J. B. & Jordan, T. H. *J. geophys. Res.* **83**, 5331–5354 (1978).
- Baranowski, J., Armbruster, J., Seeber, L. & Molnar, P. *J. geophys. Res.* **89**, 6918–6928 (1984).
- Lyon-Caen, H. & Molnar, P. *Tectonics* **4**, 513–538 (1985).
- Molnar, P. *Ann. Geophysicae* **5**, 663–670 (1987).
- Tapponnier, P. & Molnar, P. *J. geophys. Res.* **84**, 3425–3459 (1979).
- Peltzer, G., Tapponnier, P., Zhang, Z. & Xu Z.-q. *Nature* **317**, 500–505, (1985).
- Tapponnier, P., Peltzer, G. & Armijo, R. *Spec. Publ. Geol. Soc. Lond.* **19**, 115–157 (1986).
- McCaffrey, R. *J. geophys. Res.* **93**, 15,161–15,182 (1988).
- Houseman, G. A. & England, P. C. *J. geophys. Res.* **86**, 3651–3663 (1986).
- England, P. C. & Houseman, G. A. *J. geophys. Res.* **86**, 3664–3676 (1986).
- England, P. & Houseman, G. *J. geophys. Res.* (in the press).
- England, P. C. & McKenzie, D. P. *Geophys. J. R. astr. Soc.* **70**, 295–321, (1982).
- Vilotte, J. P., Daignières, M. & Madariaga, R. *J. geophys. Res.* **87**, 10709–10728 (1982).
- Vilotte, J. P., Daignières, M., Madariaga, R. & Zienkiewicz, O. C. *Phys. Earth planet. Inter.* **36**, 236–259, (1984).
- Davy, P. & Cobbald, P. *Bull. Geol. Inst. Uppsala* **14**, 129–142 (1988).
- Freund, R. *J. Geol.* **78**, 188–200 (1970).
- Garfunkel, Z. & Ron, H. *J. geophys. Res.* **90**, 8589–8602 (1985).
- Jackson, J. & McKenzie, D. *Geophys. J.* **93**, 46–73 (1988).
- Chang Cheng-fa. & Cheng Hsi-lan *Sci. Sinica* **16**, 257–265 (1973).

ACKNOWLEDGEMENTS. We thank J. Jackson and P. Tapponnier for comments. This research was supported by NERC, NASA and by a Royal Society Guest Fellowship.

Mapping the tephra fallout risk: an example from Vesuvius, Italy

F. Barberi*, G. Macedonio†, M. T. Pareschi‡
& R. Santacroce*§

* Dipartimento di Scienze della Terra, Università di Pisa, Via Santa Maria, 53-56100 Pisa, Italy

† Centro di Studio per la Geologia Strutturale e Dinamica dell'Appennino CNR, Via Santa Maria, 53-56100 Pisa, Italy

‡ Centro di Ricerca IBM, Via Santa Maria, 67-56100 Pisa, Italy

§ Istituto Internazionale di Vulcanologia CNR, Piazza Roma, 2-95123 Catania, Italy

THE goal of evaluating the risks related to the reactivation of a quiescent volcano requires the reconstructions of the eruptive history of the volcano, the construction therefrom of a behavioural model of the volcano so as to define the ‘maximum expected event’ and the subsequent quantitative models allowing reliable simulation of such an event to be set up and hazard and risk maps to be developed. We have followed this approach to infer the size and character of the explosive eruption of Vesuvius to be expected after a 45-year rest period^{1,2}, and we used a numerical model, tested on the recent Mount St Helens eruption and the Vesuvius eruption in AD 79^{3,4}, to simulate tephra transport and fallout. By combining the fallout model with a statistical analysis of the wind regime and with the density of urban settlement, we were able to assess quantitatively the tephra fallout risk.

During its long documented history, Vesuvius has exhibited very different types of activity ranging from quiet lava emission to moderate strombolian explosions, to catastrophic, highly

explosive (plinian) eruptions. Over the past 17,000 years, there has been a positive correlation between the length of the repose interval and (1) the size and explosiveness of the following eruption, and (2) the degree of evolution of erupted magma. This behaviour is considered to be the consequence of nearly continuous supply of deep basic magma to a shallow (3–5 km depth) magma chamber, in which differentiation and mixing occur. Since AD 79 (the year of the famous 'Pompei' eruption), the erupted products have all belonged to the 'high-potassium' series and the magma supply rate has been roughly constant, $1.5\text{--}2.0 \times 10^6 \text{ m}^3 \text{ yr}^{-1}$ (ref. 2). Assuming that no large changes have occurred in the magmatic system since the most recent eruption (March 1944), a volume of $50\text{--}100 \times 10^6 \text{ m}^3$ of magma could now be available in the shallow reservoir of the volcano. Although only an estimate, a large error in this volume seems quite unlikely, based on the history of Vesuvius.

The emission of such a volume of magma during a single eruption would be second only to the severe subplinian eruption in 1631. The most probable type of reactivation is an explosive subplinian eruption with the usual sequence of events (pyroclastic fall, pyroclastic flows and mud flows) that occur at Vesuvius during this kind of event. This can be considered the 'maximum expected event' and we have previously^{5,6} simulated it numerically. Here we attempt a quantitative assessment of the risk zonation for tephra fallout from such a subplinian eruption.

Besides the volcanological input data (summarized in Table 1), we needed a statistical study of the present wind regime in southern Italy. Data on the wind velocity and direction at different altitudes during 1962–1976 are available from the Brindisi meteorological station. We selected 3,125 wind profiles (between zero and 32,000 m above sea level; Fig. 1) for use in our model of fallout from the eruption column. The simplified

model that we used is based on the solution of a three-dimensional transport equation for tephra particles. The equation accounts for the advection by wind, for the settling of the erupted particles by gravity, and for atmospheric turbulent diffusion. The space is divided into vertical steps of 2 km, and the eruptive column is approximated as a vertical-line source that expands by turbulent diffusion. The vertical distribution of mass inside the column and the maximum column height are calculated according to previous models^{3–9}.

The calculated position of the centre of mass of the deposit for each wind profile allows us to determine the percent occurrence of the direction of the main dispersal axis of the deposit. The number of wind profiles considered (3,125) is large enough for a meaningful statistical analysis with the chosen 15° tolerance in the direction of the dispersal axis: by assuming a Poisson distribution inside each direction interval, the error ranges between 5.5% and 20% with a mean error of 12%. We found no significant differences related to seasonal wind regimes, a fact that reflects the prevailing high-velocity western winds operating between altitudes of 8 and 15 km during the whole year. Over 18–20 km the summer winds are clearly distinguished from those of other seasons by their prevailing eastern provenance¹⁰. But even for eruption columns much higher than those considered in this work, the low velocity which characterizes these winds ($\sim 10 \text{ m s}^{-1}$) would not significantly affect the main dispersal axis.

Once the deposit is calculated, the area with ground concentration higher than a previously determined threshold Th is identified. The procedure is repeated for each of the 3,125 available wind profiles, and the collection of areas with above-threshold fallout maps the probability for above-threshold surface tephra concentration. On the ground a grid of points

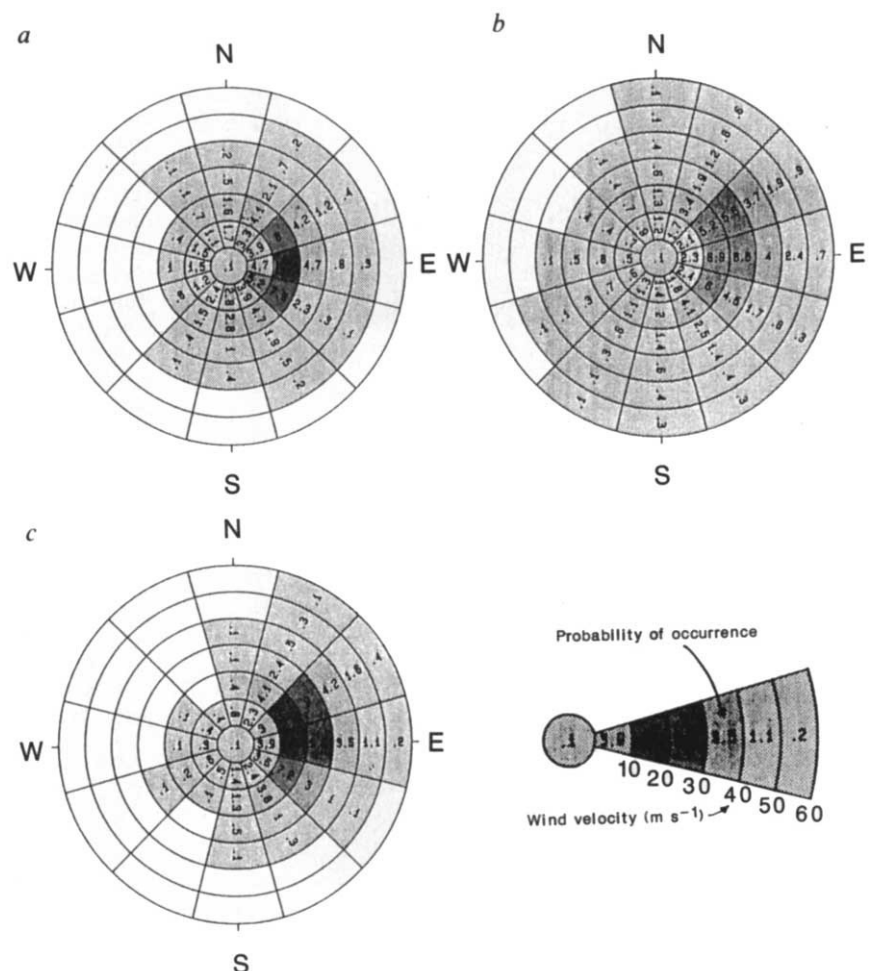


FIG. 1 Polar diagram showing the wind velocity distribution in southern Italy during 1962–1976 (3,125 data from Brindisi meteorological station) at altitudes of a, 5 km; b, 9 km; c, 13 km.

TABLE 1 Volcanological data

Eruptible mass	2×10^{11} kg
Mass eruption rate	10^7 kg s $^{-1}$
Exsolved gas	5–10 wt%
Diffusion coefficients*	3,000 m 2 s $^{-1}$
Maximum column height	13 km
Temperature	1,000 °C

Size and density of particles are found from previous eruptions of Vesuvius. The particle settling velocities are taken from ref. 6.

* Vertical diffusion is neglected.

600×600 m 2 was chosen. The ratio between the number of wind profiles giving a concentration higher than Th and the number of available wind profiles gives the probability of a coverage heavier than Th at that point. These data can therefore be used to draw a hazard map related to the 'maximum expected event', where hazard is identified with the probability that a ground deposition exceeds a given threshold. In Fig. 2 two cases are presented (Th = 100 and 200 kg m $^{-2}$). The extent of the area expected to be covered by >100 (200) kg m $^{-2}$ ranges from ~ 10 km 2 to more than 1,000 (800) km 2 for a probability decreasing from 60% (45%) to 2% (1%). We tested the sensitivity of

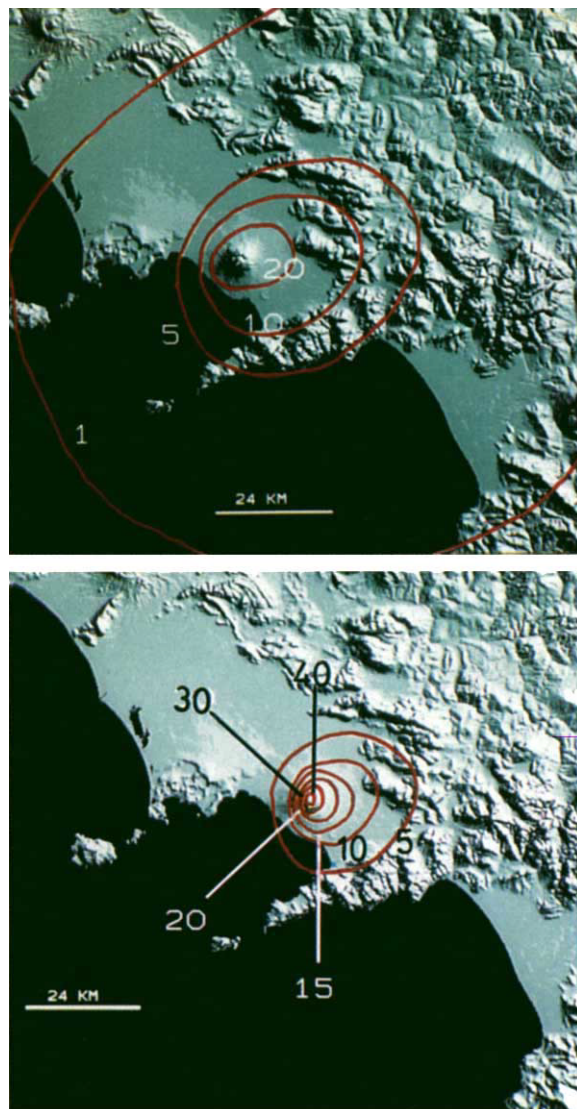


FIG. 2 Hazard maps related to tephra fallout in the Vesuvius area. Lines define areas of isoprobability of having tephra blankets heavier than (top) 100 kg m $^{-2}$ and (bottom) 200 kg m $^{-2}$.

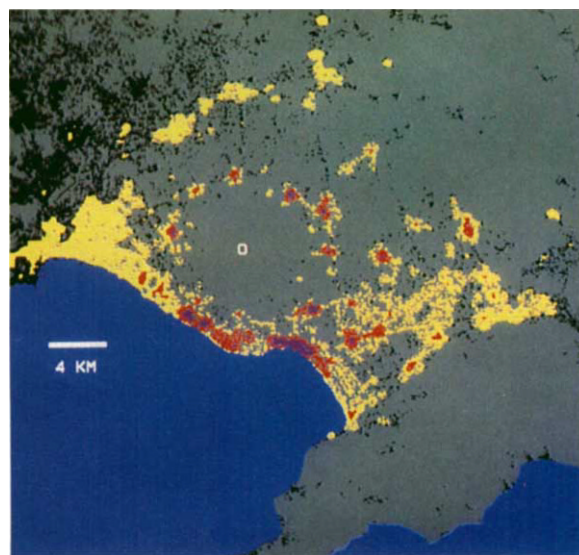


FIG. 3 Roof-collapse risk map related to tephra fallout accumulation heavier than 200 kg m $^{-2}$ in the Vesuvius area. The risk scale (between 1 and 100) represents the product (hazard $\times$ value), vulnerability being equal to unity. Grey = 0; black = 1–5; yellow = 5–10; red = 10–20; violet = 20–100.

the results to column height variations: the probability, in a given area, of a ground concentration >100 and 200 kg m $^{-2}$ increases by a few per cent as the column height decreases in the range 18–9 km. This is the result of the assumption of constant erupted mass: a column height increase would therefore result into higher eruption rate, lower eruption duration and wider dispersal of tephra.

The hazard maps (assessed for the 'maximum expected event') provide the basis for the risk zonation for tephra fallout. According to the Working Group on the Statistical Study of Natural Hazards¹¹, risk is the possibility of loss. From a quantitative point of view, it can be defined by the product: value $\times$ vulnerability $\times$ hazard. Value can represent the number of lives or capital value (buildings, for example) or productive capacity exposed to hazard and can depend, in the case of tephra fallout, on several factors, such as the accumulation of pyroclastic fall deposits that may cause failure and collapse of roofs, the deposition of ash that may destroy vegetation and impede road traffic. Vulnerability is that fraction of value that is likely to be lost as a result of the event.

As an application we present a risk map of the Vesuvius area for roof collapse under the load of fall deposit (Fig. 3). The density of urban settlement (moving average of inhabited areas on 600×600 m 2 areas from a 1984 Landsat TM image) has been assumed to be representative of the 'value' and a weight greater than 200 kg m $^{-2}$ has been considered enough to cause collapse of 100% of the roofs (vulnerability = 1). We map the risk by plotting, at each point, the product of the 'value' and the probability of having more than 200 kg m $^{-2}$ of pyroclastic fall deposit. In Fig. 3 colours from black to yellow, red and violet emphasize the urban areas with increasing risk. Civil-defence preparedness plans should be concentrated in these areas. □

Received 12 October 1989; accepted 22 January 1990.

1. Santacroce, R. (ed.) *Somma-Vesuvius*, Rep. No. 114-8 (CNR, Rome 1987).
2. Santacroce, R. *J. Volcanol. geotherm. Res.* **17**, 237–248 (1973).
3. Armienti, P., Macedonio, G. & Pareschi, M. T. *J. geophys. Res.* **93**, 6463–6476 (1988).
4. Macedonio, G., Pareschi, M. T. & Santacroce, R. *J. geophys. Res.* **93**, 14817–14827 (1988).
5. Macedonio, G., Pareschi, M. T. & Santacroce, R. *Proc. Int. Conf. on Volcanoes 487–490* (National Institution for Research Advancement, Tokyo, 1988).
6. Macedonio, G., Pareschi, M. T. & Santacroce, R. *J. Volcanol. geotherm. Res.* (in the press).
7. Sparks, R. S. J. *Bull. Volcanol.* **48**, 3–16 (1986).
8. Woods, A. W. *Bull. Volcanol.* **50**, 169–193 (1988).
9. Wilson, L. & Walker, G. P. L. *Geophys. J. R. astr. Soc.* **89**, 657–679 (1987).
10. Cornell, W., Carey, S. & Sigurdsson, H. *J. Volcanol. geotherm. Res.* **17**, 89–109 (1983).
11. *Report of Consultative Meeting of Experts on the Statistical Study of Natural Hazards and their Consequences* Doc. SC/WS/500, 1–11 (UNESCO, Paris, 1972).

Early Carboniferous low-temperature hydrothermal vent communities from Newfoundland

Peter H. von Bitter^{*†}, Steven D. Scott[†]
& Paul E. Schenk[‡]

^{*} Department of Invertebrate Palaeontology, Royal Ontario Museum, Toronto M5S 2C6, Canada

[†] Department of Geology, University of Toronto, Toronto M5S 3B1, Canada

[‡] Department of Geology, Dalhousie University, Halifax B3H 3J5, Canada

LIVING vent communities, dependent on microbial chemosynthesis rather than photosynthesis, were discovered only in the 1970s¹. Since then, six fossil vent communities have been identified; they range in age from early Carboniferous to early Tertiary²⁻⁷. Here we describe fossil tubes, an abundant low-diversity fauna and sulphide mineralization in 340-Myr-old Carboniferous carbonate mounds in Newfoundland (Fig. 1a). These features, together with evidence for microbial activity, point to the existence of a seventh chemosynthetic community, clustered around low-temperature hydrothermal vents. The remarkable preservation of this New-

foundland community allows a direct comparison to be made with the composition and structure of modern vent communities.

We observed the fossil tubes on the Port au Port Peninsula (Fig. 1b) in mounds of the Big Cove Formation, Lower Codroy Group, which unconformably overlie Ordovician carbonates^{8,9} (Fig. 1c). Clusters of tubes occur in apparent growth position in Mistaken Cove (Fig. 2). The concentration of fossil tubes in the lowest 2 m of the larger mounds may be simply due to inaccessibility of the higher points on the mounds.

Fossil-tube fragments, generally encrusted by microbial colonies, are up to 20 cm long (Fig. 2), 1.5–3 cm in diameter, circular to elliptical in cross-section, depending on lateral compression (Fig. 3a), taper longitudinally, and commonly curve slightly. The rare apex lacks evidence of a holdfast. Macroscopically, tube walls appear laminated; microscopically, the layers are diagenetically altered (Fig. 3b). External micro-ornamentation and internal structures were not observed. Several generations of sub-parallel geopetal infill are usually present (Fig. 3a, b). Juvenile tubes were not recognized.

The carbonate mounds vary in size, reaching 10 × 16 m in cross-section. They consist of ~40% original void space, filled with several generations of calcite cement and internal sediment. The mound framework consists mainly of exceptionally well developed thickets of trepostome bryozoans (*Stenoporella*?),

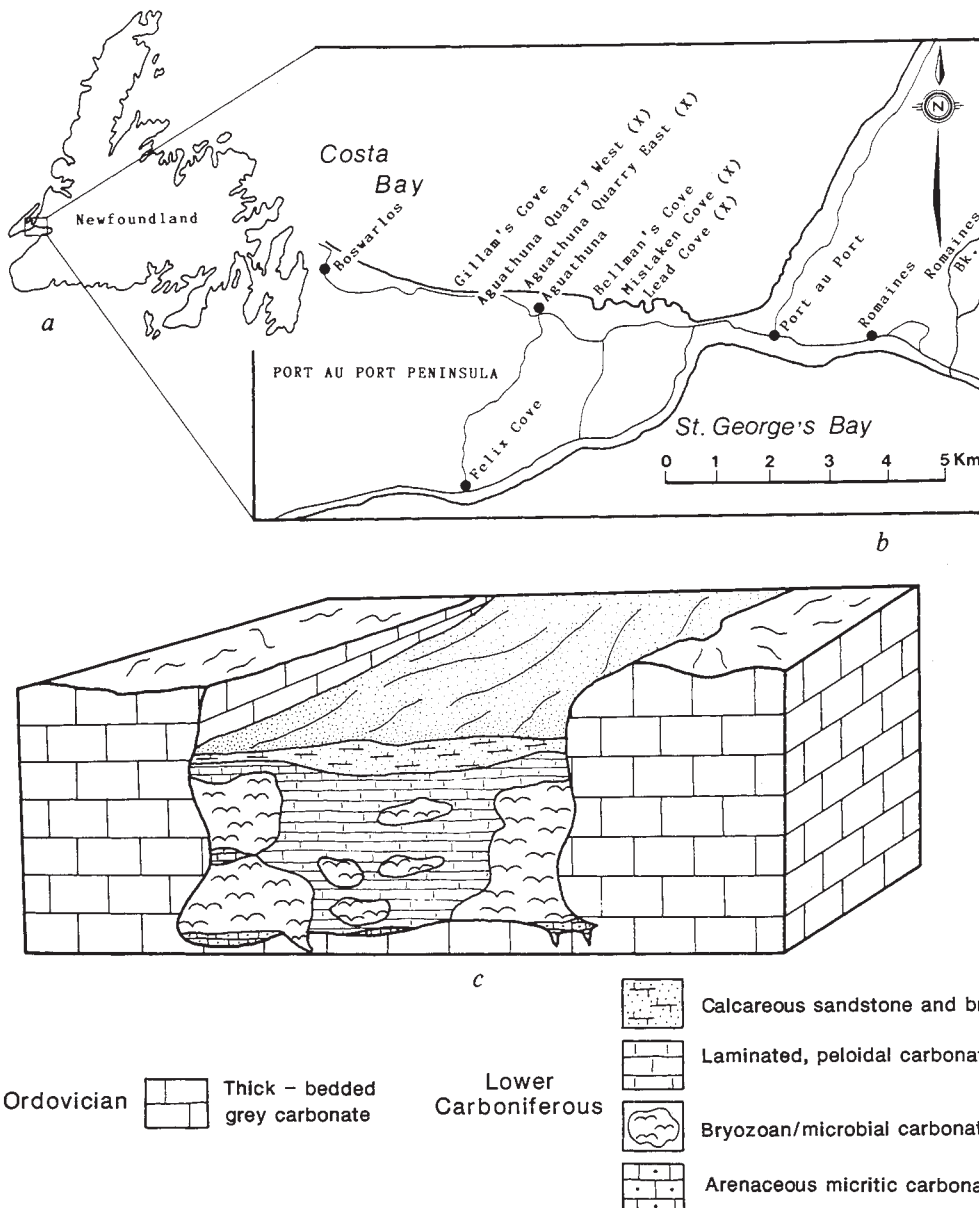


FIG. 1 a, Port au Port Peninsula, Newfoundland, Canada. b, Fossil-tube localities marked by X. c, Schematic cross-section of lower stratigraphic units of the Big Cove Formation in the Bellman's Cove area infilling a karst valley in Ordovician carbonates. Overlying sandstones and conglomerates not shown. Maximum depth and width of karst valleys are ~16 m and 300 m, respectively. Modified from ref. 8.

encrusted with microbial coatings⁸. Concentrations of unusually well preserved brachiopods (*Beecheria*) are locally present. Laminated carbonates and peloidal shaley carbonates are inter-mound equivalents of the mounds (Fig. 1c). They are overlain at Romaines Brook and Boswarlos (Fig. 1b) by thick marine evaporites⁹.

Metallic sulphide mineralization of spongy marcasite and euhedral pyrite, iron-poor sphalerite and galena, is abundant in the tube-bearing mounds. Barium and strontium sulphates and carbonates are intergrown with the sulphides. Some large masses of colloform and euhedral sulphides are mineralogically and texturally similar to Mississippi Valley-type ores¹⁰. The first three localities east of Aguathuna (Fig. 1b) have gossans of weathered sulphides. Vertical to sub-vertical centimetre-wide calcite veins with minor galena, marcasite and barite cut Ordovician carbonate beneath the mounds at Gillam's Cove. Sub-vertical, pipelike, mineralized cavities up to 30 cm wide occur in the mounds at Mistaken Cove.

Our hypothesis of Lower Carboniferous vent communities in Newfoundland is based on several coinciding geological observations and on comparison with modern vent communities. We use the word 'vent' generically, meaning an opening for the escape of fluid, without implying any particular rates, temperature of discharge or water depth. Although initial discoveries of vents with tubeworms and sulphides were in the deep sea, chemosynthetic communities can occur in a variety of marine settings. Availability of a nutrient source (hydrogen sulphide and/or methane) for chemosynthesis, rather than whether that source is 'hot', 'cold', 'deep' or 'shallow', determines the presence or absence of such communities¹¹. Chemosynthetic communities are found in low-temperature settings at various water depths: in <100 to ~150 m fossil deposits^{4,7}, at 500–900-m hydrocarbon vents¹², at 3,266-m cold sulphide seeps¹³, and at 3,800–3,900-m hydrocarbon vents¹⁴.

Similarity exists between presently forming carbonate mounds, with their excessive biological growth (including tubeworms), and disseminated sulphides in the Guaymas Basin, Gulf of California^{15,16}, and the features described from Newfoundland; however, carbonate from the Guaymas Basin is primarily inorganic, whereas in Newfoundland it is mostly organic. Bacterial activity, as evidenced by extensive bacterial mats at Guaymas Basin^{11,17}, may be comparable to that which formed the laminated and peloidal intermound carbonates. Guaymas Basin is a deep (2,000 m) sea-floor spreading site with

'black smoker' vents up to 359 °C (ref. 15); water depth, and mineralization temperatures were much less in the Newfoundland deposits, with their strike-slip faulting, rift-valley tectonic setting¹⁸. Closer modern analogues may be the carbonates being formed, and associated abundant tubeworms and clams being fed, by methane-enriched fluids expelled from accretionary sediments along the Oregon–Washington margin¹⁹ and from a transform fault zone at the foot of the continental slope along the northern margin of Guaymas Basin²⁰.

Although the presence of large and abundant fossil tubes in Newfoundland is explainable by a vent hypothesis, their identity needs confirmation. Tube-forming organisms are disparate and



FIG. 2 Sulphide-associated, 340-million-year-old fossil tubes in vent-formed bryozoan/microbial carbonate mound, Mistaken Cove, western Newfoundland, Canada. In possible growth position and underlain by black carbonaceous shale which may have been a source of H_2S and CH_4 for chemosynthetic bacteria. Diameter of coin (white oval) is 2.35 cm.

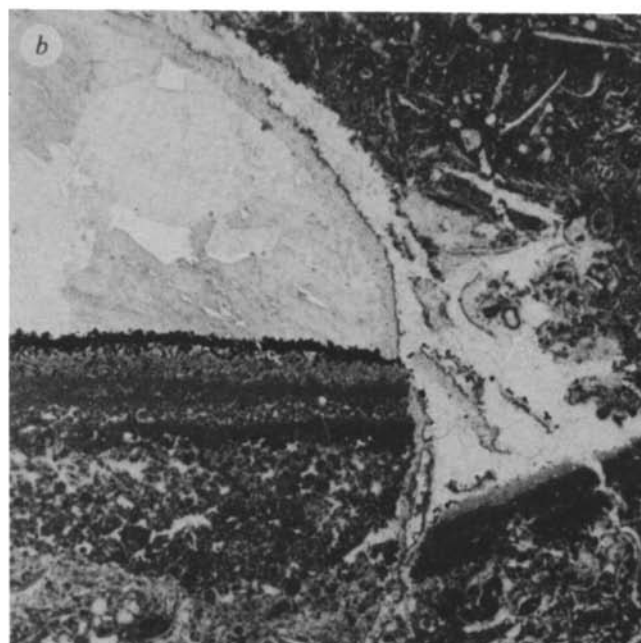
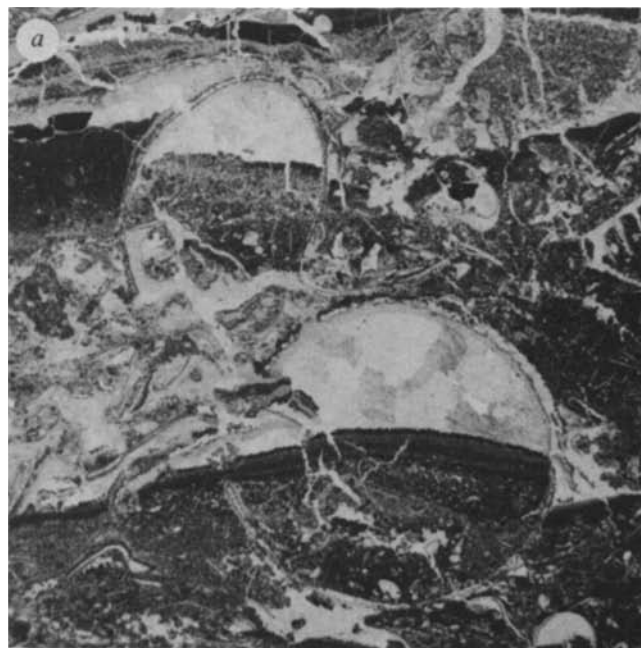


FIG. 3 a, b, Transverse sections of fossil tubes in bryozoan/microbial carbonate mound. Note double tube-wall structure, several generations of cavity infill and uppermost dark black horizontal layer of sphalerite. Mistaken Cove, Newfoundland, ROM 45879, crossed nicols, magnified 2.2 and 4.7 times, respectively.

include segmented bristleworms, pogonophorans and vestimentiferans (the latter commonly associated with deep-sea vents). The size, general morphology and geological context of the fossils suggests comparison with vestimentiferan tubeworms, such as the thick-walled *Riftia*¹⁷, that live in chitinous tubes up to 1.5 m long with an average diameter of 4 cm (ref. 21). Four of the five vestimentiferan families have double-wall structure and tapered tubes comparable to the laminated wall of the fossil tubes (Fig. 3b) and the pointed apex of some fossil tubes. The outer surfaces of many living Vestimentifera bear faint to pronounced annulations and/or flanges; these have not been observed on the fossils owing to their delicate nature, the loss of chitin during fossilization, encrustation, or because some specimens are internal moulds.

The uncommon, low-diversity composition of the mound biota is directly explainable by a vent origin. The mounds contain bryozoa, brachiopods, the tubes, conularids, serpulid worms and crustaceans, but lack normal Carboniferous groups such as corals, echinoderms and macrocalcereous algae⁸. This is similar to low-diversity faunas lacking corals, sponges and bryozoa, with rare echinoderms, that characterize modern vent communities¹⁷. The lack of diversity in both ancient and modern vent environments is attributable to the fact that only "certain specifically adapted invertebrates can live in the specialized vent environment"¹¹. The Newfoundland biotas have been regarded as 'stressed'⁸ which we interpret to be due to both schizohaline conditions⁸ and living beside vents.

Extant vent communities share remarkable abundance and unequal distribution with the Newfoundland mound biotas. Abundant, well-preserved, closed valves of the brachiopod *Beecheria* commonly occur as 'nests' up to 0.5 m² in area. The colonies are in shell support, often with some sediment between the valves; the shells have a normal size-frequency distribution⁸.

The above suggests that the mounds resulted from colonization by sessile, filter-feeding organisms (Fig. 4) that were either in a symbiotic relationship²² with chemosynthetic bacteria or fed directly on them²³. Tubeworms in this environment became encrusted, the same as living vent-associated species²⁴. The sub-parallel, multigenerational, infilling peloidal sediments (Fig. 3a, b) were probably produced within the tubes by successive generations of bacteria, between episodes of bacteria-induced precipitation of spar calcite. The dense colonies of the filter-feeding *Beecheria* are explained by colonization and feeding at a superabundant food supply. The shell-supported brachiopod colonies, their unusually perfect preservation, and the range of juvenile to mature individuals all suggest *in situ* growth with little or no transport after death. The bryozoan/microbial thickets and abundant evidence for bacterial activity suggest a possible symbiotic relationship between the bryozoa and bacteria. Galatheid and brachyuran crabs and filter-feeding serpulid worms at extant vent sites¹⁷ may be analogues of the arthropods and serpulid worms reported^{8,9} from Newfoundland.

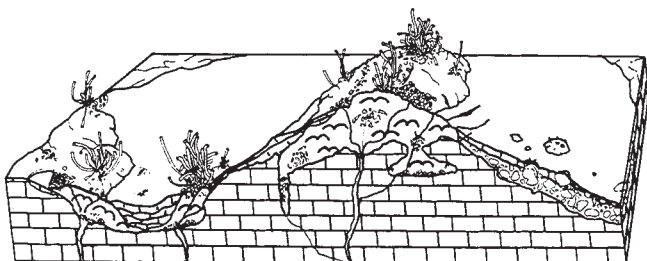


FIG. 4 Early Carboniferous chemosynthetic community. Carbonate mound formation at sites of concentrated venting; peloidal and/or laminitic carbonate formation at sites of diffused venting. Lead Cove/Boswarlos area, Newfoundland. Lithological symbols same as in Fig. 1c, except that coarse conglomerates (on right) have been added. Lateral scale condensed (see Fig. 1b).

A necessary requirement in support of chemosynthesis is for the sulphides to have been emplaced during, rather than substantially after, mound growth. Strong evidence for this synchronicity is provided by the relationship of mineralized veins and the host rock, by local and regional concentrations of mineralization in the stratigraphically lowest Lower Carboniferous marine carbonates, and by its near exclusion in other units. Mineralizing solutions seem to have accessed the mounds through sulphide-containing calcite veins cutting the underlying Ordovician limestone. At Gillam's Cove, fractures in the latter contain fossiliferous Carboniferous sediment which is, in turn, cut by mineralization. The mounds are not veined and fluid flow in them was diffuse. Codroy Group mineralization is concentrated in the stratigraphically lowest carbonate mounds, and at the contact between the laminated and peloidal intermound carbonates and underlying units. Significant mineralization is apparently lacking within the intermound facies or in the thick marine sequence that overlies it in some places. These relationships suggest that the sulphide mineralization was introduced during the early Carboniferous, simultaneous with mound formation.

The relationship of sulphides to mound sediments and fossils is less definitive. Most iron sulphide was introduced with relatively late cavity-filling calcite, but the timing of Pb/Zn mineralization is uncertain. Euhedral crystals of sphalerite, both as the youngest layer of internally discordant geopetal infilling (Fig. 3a, b) and as possible lag deposits in mound cavities suggest that different generations of sediment were infilling loose tubes synchronously during mound growth; alternatively, the discordancy may represent cross-bedding caused by later carrying currents. Radiometric isotope studies are similarly equivocal. The Newfoundland mounds are Middle Viséan (~340 Myr), based on comparison with faunas from similar Nova Scotian mounds²⁵. Galena from them yields model lead ages ranging from Carboniferous to Permian (308–252 Myr)²⁶.

The temperature of sulphide mineralization can be inferred from homogenization temperatures of fluid inclusions in the associated calcite and demonstrates that the mineralization is truly hydrothermal. All primary fluid inclusions examined are liquid vapour type, without daughter minerals, and show no evidence²⁷ of having been trapped under boiling conditions. Mound material from Lead Cove gave uncorrected homogenization temperatures of 65 to 187 °C, with most between 80 and 120 °C. Vein calcite from beneath the mounds at Gillam's Cove gave a similar range of 73–178 °C. Salinities determined by freezing-point depression range from 15 to 21 equivalent wt% NaCl. The pressure correction to be added to the homogenization temperatures, for the relatively shallow marine setting envisaged for the Newfoundland mounds, is probably no more than ~10 °C (ref. 27). The highest-temperature fluid inclusions require a water depth of at least 100 m for these fluids not to have boiled. Conodont colours from the mound and laminite facies indicate that the host rocks were never at temperatures above 90 °C. The fluid inclusion and colour data are consistent with the sulphides having been deposited by low-temperature, saline hydrothermal fluids not unlike those responsible for Mississippi Valley-type mineralization^{10,28}. □

Received 9 October 1989; accepted 26 January 1990.

- Corliss, J. B. *et al. Science* **203**, 1073–1083 (1979).
- Haymond, R. A., Koski, R. A. & Sinclair, C. *Science* **223**, 1407–1409 (1984).
- Oudin, E. & Constantinou, G. *Nature* **308**, 349–353 (1984).
- Banks, D. A. *Nature* **313**, 128–131 (1985).
- Gaillard, C. *et al. Bull. Soc. géol.* **9**, 69–78 (1985).
- Oudin, E., Bouladon, J. & Paris, J.-P. *C. R. Acad. Sci. Paris* **301**, 157–162 (1985).
- Beauchamp, B. *et al. Science* **244**, 53–56 (1989).
- Dix, G. R. & James, N. P. *Sedimentology* **34**, 779–793 (1987).
- Dix, G. R. & James, N. P. *Can. J. Earth Sci.* **26**, 1089–1100 (1989).
- Anderson, G. M. & MacQueen, R. W. *Geoscience Can.* **9**, 108–117 (1982).
- Jannasch, W. & Mottl, M. J. *Science* **229**, 717–725 (1985).
- Brooks, J. M. *et al. Eos* **68**, 498–499 (1987).
- Paull, C. K. *et al. Science* **220**, 965–967 (1984).
- Mayer, L. A. *et al. Deep Sea Res.* **35**, 1235–1246 (1988).

15. Lonsdale, P. & Becker, K. *Earth planet. Sci. Lett.* **73**, 211–225 (1985).
16. Peter, J. M. & Scott, S. D. *Can. Miner.* **26**, 567–587 (1988).
17. Grassle, J. F. *Science* **229**, 713–717 (1985).
18. Bradley, D. C. *Tectonics* **1**, 107 (1982).
19. Ritger, S., Carson, R. & Suess, E. *Geol. Soc. Am. Bull.* **98**, 147–156 (1987).
20. Simoneit, B. R. T. et al. *Appl. Geochem.* (in the press).
21. Jones, M. L. *Science* **213**, 333–336 (1981).
22. Childress, J. J., Felbeck, H. & Somero, G. N. *Scient. Am.* **256**, 115–120 (1987).
23. Stein, J. L. *Science* **223**, 696–698 (1987).
24. Tunnicliffe, V. & Fontaine, A. R. *J. geophys. Res.* **92**, 11,303–11,314 (1987).
25. Giles, P. S., Boehner, R. C. & Ryan, R. J. *N.S. Dept Mines Pap.* **79**–7, 1–57 (1979).
26. Swinden, H. S., Lane, T. E. & Thorpe, R. I. *Can. J. Earth Sci.* **25**, 593–602 (1988).
27. Roedder, E. *Miner. Soc. Amer.* **12**, 1–646 (1984).
28. Hanor, J. S. in *Geochemistry of Hydrothermal Ore Deposits* 2nd edn (ed. Barnes, H. L.) 137–172 (Wiley, New York, 1979).

ACKNOWLEDGEMENTS. We thank B. Boyle, J. Burke, K. Coates, G. Dix, L. Mayer, R. Matsumoto, A. McCol, J. Peter, D. Rudkin, B. Simoneit, E. Spooner, S. Swinden, and R. von Bitter for information, assistance and discussion.

Evolutionary genetics of host use in swallowtail butterflies

John N. Thompson, Wayne Wehling
& Robert Podolsky *

Departments of Botany and Zoology, Washington State University,
Pullman, Washington 99164, USA

* Present address: Department of Botany and Plant Sciences,
University of California, Riverside, California 92521, USA

PHYTOPHAGOUS insects include many of the most species-rich genera and families worldwide^{1,2}, and much of that diversity seems to result from speciation onto different plant species^{3–7}. Divergence onto different host plants is thought to involve at least two sets of loci, which may be genetically linked: loci controlling the preference of ovipositing females for a particular plant species, and loci controlling the ability of larvae or nymphs to feed on those plants^{8–11}. Others have argued that oviposition preference and larval performance may be pleiotropic effects of the same loci^{12–15}. The genetic relationship between oviposition preference and larval performance has therefore become a central problem in the developing theory of insect and plant interactions^{16–23} and speciation in insects^{24–27}. Results from interspecific crosses between two swallowtail butterfly species that feed on different plant families indicated that oviposition preference is controlled in these insects primarily by one or more loci on the X chromosome²⁸. In a series of reciprocal interspecific crosses between these species,

Papilio zelicaon and *Papilio oregonius*, we investigated whether larval performance on different plant species was also controlled by X-linked loci, which could allow for strong correlations between oviposition preference and larval performance. We found no X-chromosome effect for any component of larval performance. Loci from both parents influenced survivorship, and maternal effects influenced pupal mass and possibly development time. The results varied with the specific measure of performance, the host plant species used in comparing reciprocal crosses, and the sex of the larvae, all of which caution against the use of any single measure of performance in evaluating the evolutionary genetics of host shifts.

About 560 species of swallowtail butterflies (family Papilionidae) are known worldwide²⁹, and at least 50 plant families have been recorded as hosts for the larvae³⁰. This diversification onto such a wide variety of plant taxa has made this group a focus of research on the evolution of interactions between insects and plants^{31–36}. *P. zelicaon* and *P. oregonius* are two closely related swallowtail species from western North America that differ in the plant families on which their larvae feed. *P. zelicaon* larvae feed mostly on plants within several genera of the Umbelliferae, whereas *P. oregonius* larvae feed exclusively on *Artemisia dracunculoides* (Compositae) in the field. These butterfly species are sufficiently closely related, however, that viable and fertile interspecific hybrids can be obtained by individually hand-pairing the adults, providing an opportunity to analyse the genetics of host use.

Full-sib families of *P. zelicaon* and *P. oregonius* were initially tested for their ability to feed on *Lomatium grayi*, a common umbelliferous host for *P. zelicaon*, and *A. dracunculoides*, the only host for *P. oregonius* in the field. *P. zelicaon* families were second generation offspring of wild-caught females collected at Sailor Bar, near Sacramento, California; *P. oregonius* families were second generation offspring of wild-caught females collected at Palouse Falls, Franklin County, Washington. The parents of all larvae used in the experiments had been reared on *Foeniculum vulgare* to mitigate any maternal effects on offspring that might result from differences in the plant species on which the parents had been reared. This introduced plant species is highly suitable for development of both *P. zelicaon* and *P. oregonius*. In all, 428 larvae were individually monitored throughout development in these initial trials. Rearing procedures are given in Fig. 1.

Two series of interspecific reciprocal crosses were also made using these populations. In series 1, six offspring (3 females and

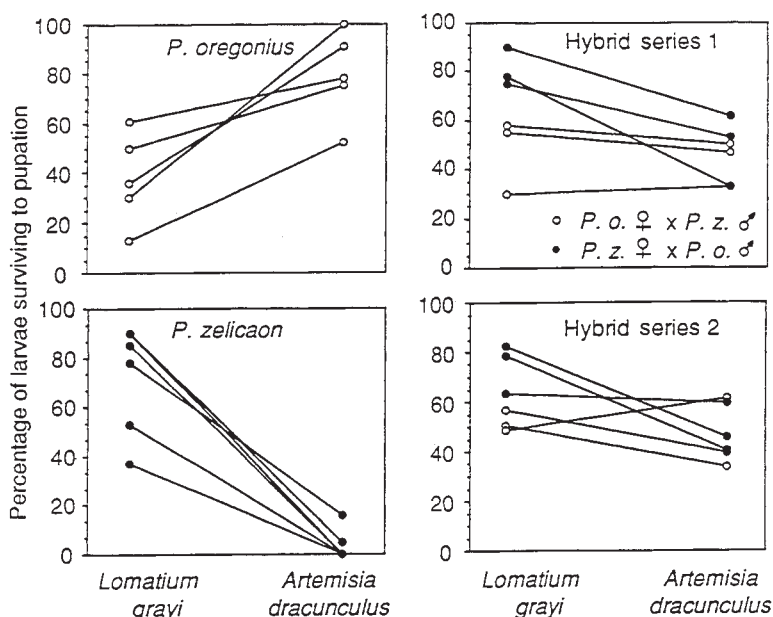
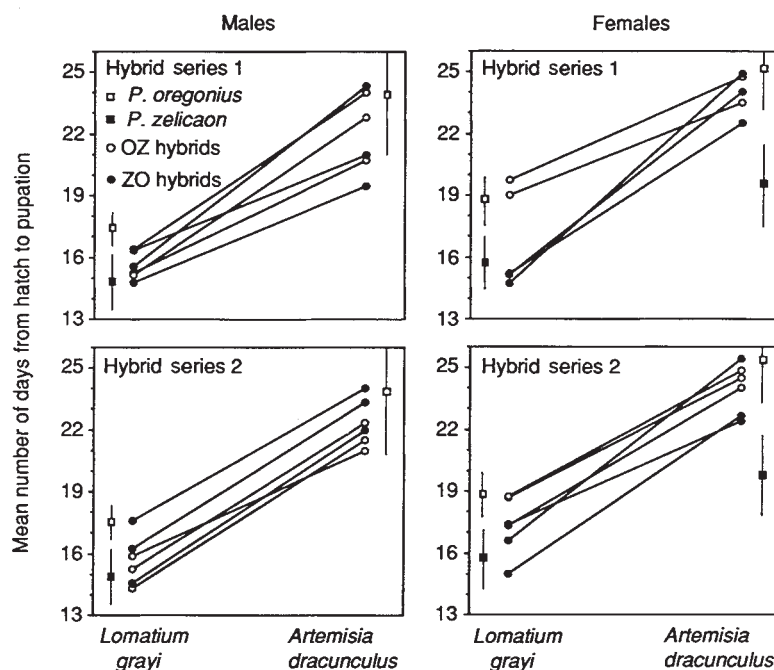


FIG. 1 Survivorship of *Papilio* spp. and interspecific hybrid larvae from egg hatch to pupation on two plant species in different plant families. For each trial, 36–44 (except for one trial on *P. oregonius* with $n=59$ and another with $n=16$) newly hatched larvae from each female were split equally onto vegetative sprigs of the two plant species. Each sprig had its base immersed in a florist's waterpick (a device to prevent wilting). Initially, two larvae were placed on each sprig, which was then put into a 150 × 25 mm Petri dish. After the second instar, all larvae were divided into one per dish. Larvae were kept in an environmental chamber set at 23 °C, and spent 16 h in the light and 8 h in the dark. Each larva was checked daily, and food was changed every 2 d or sooner. The position of the Petri dishes within the environmental chamber was changed daily to minimize position effects.

FIG. 2 Mean number of days for interspecific hybrid larvae to develop from egg hatch to pupation. OZ=*P. oregonius* female $\times$ *P. zelicaon* male; ZO=*P. zelicaon* female $\times$ *P. oregonius* male. Means ($\pm$ s.d.) are also shown for non-hybrid *P. oregonius* and *P. zelicaon*, with all families combined within species, as a basis for comparison with hybrids. Only one male *P. zelicaon* reached pupation on *A. dracunculus* (20 d), so no mean is shown on the figure.



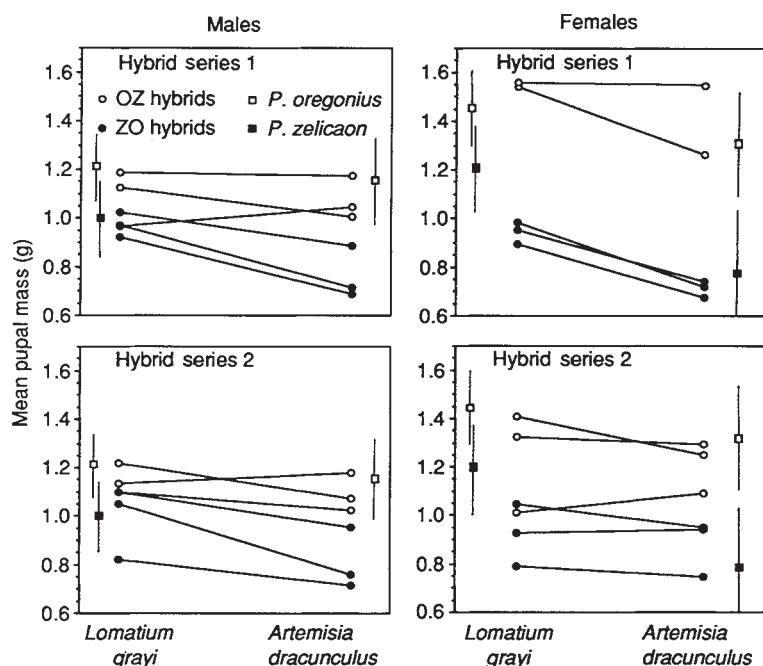
3 males from family 952) of one *P. zelicaon* female were hand-paired singly with six offspring (from family 931) of one *P. oregonius* female to produce 6 interspecific families. In series 2, six offspring (3 females and 3 males from family 953) of a different *P. zelicaon* female were hand-paired singly with six other offspring of the same *P. oregonius* female used in series 1 (family 931). All individuals used in the matings had been reared on *F. vulgare*. The procedures for rearing hybrid larvae were the same as for the non-hybrid larvae. In all, 470 interspecific larvae were monitored throughout development in these two series of crosses.

All *P. zelicaon* families show high to moderate survivorship on *L. grayi*, but very low survivorship (0–16%) on *A. dracunculus* (Fig. 1). By contrast, all *P. oregonius* families show higher survivorship on *A. dracunculus* than on *L. grayi* (Fig. 1). Hybrids are able to survive on both hosts, regardless of the direction of the cross (Fig. 1). This result indicates that no single dominant allele from one of the parental species determines survivorship

on these plant species. Instead, hybrid survivorship is influenced by loci of both parental species. Overall survivorship for both OZ (*P. oregonius* female $\times$ *P. zelicaon* male) and ZO hybrids is significantly higher on *L. grayi* than on *A. dracunculus* in both series 1 and 2 ($P=0.037$, two-way analysis of variance (ANOVA) on plant species and direction of cross, logit transformation), although in some families survivorship is roughly equal on the two hosts. The direction of the cross significantly influences survivorship in series 1 ($P=0.028$), but not in series 2 ($P=0.065$). There is no significant interaction effect between plant species and direction of the cross in either series. Survivorship of hybrid larvae, however, cannot be used to assess sex-linkage of larval performance. The sex of larvae cannot be determined and differential survival of sexes cannot be eliminated to explain how direction of the cross affects survival on the different hosts.

For non-hybrid larvae that survive to pupation, both males and females of *P. oregonius* reach pupation faster on *L. grayi*

FIG. 3 Mean mass (g) of pupae in reciprocal interspecific crosses between *P. oregonius* and *P. zelicaon*. All pupae were weighed to the nearest 0.0001 g on the day after pupation. Means ($\pm$ s.d.) are also shown for non-hybrid *P. oregonius* and *P. zelicaon*, with all families combined within species, as a basis for comparison with hybrids. Only one male *P. zelicaon* reached pupation on *A. dracunculus* (1.0 g), so no mean is shown on the figure.



than on *A. dracunculus* ($P < 0.001$, ANOVA) (Fig. 2). When reared on *L. grayi*, *P. zelacon* larvae develop on average 3 days faster than *P. oregonius* larvae. All but four of the 119 *P. zelacon* larvae reared on *A. dracunculus* died during larval development or pupation. Hybrid larvae take fewer days to reach pupation on *L. grayi* than on *A. dracunculus* ($P < 0.0001$ for both series 1 and 2, ANOVA, log transformation) (Fig. 2). There is no indication that development time is X-linked in hybrid females. Females are the heterogametic sex in Lepidoptera³⁷, so a significant effect of the X chromosome would be visible in F1 hybrids as traits resembling the paternal species. If anything, hybrid females reared on *L. grayi* are more similar to their maternal species in the number of days they take to reach pupation (Fig. 2), but there is no clear overall significant difference in development time as a result of the direction of the cross ($0.05 < P < 0.10$, ANOVA using corrected expected mean squares for mixed model). If additional experiments were to show that these differences are significant, it would indicate that development time in females is partially Y-linked, or is influenced by non-genetic maternal effects. The important point with respect to the relationship between oviposition preference and larval performance is that our results rule out X-linkage for this component of larval performance in females. Hybrid males provided no evidence of X-linkage or maternal effects. But it is more difficult to assess male hybrids than females for X-linkage because of potential complications with dominance and penetrance for X-linked loci in this sex. The overall complex effect of direction of the cross, including differences between sexes and hosts, is indicated in the ANOVA as a significant interaction effect between host species and sex of the larva in both series ($P < 0.04$), and as a significant three-way interaction between direction of the cross, hostplant and sex in series 1 ($P < 0.004$).

Pupal masses of *P. oregonius* are higher on *L. grayi* than on *A. dracunculus* and are higher in females than in males (both $P < 0.01$, ANOVA) (Fig. 3). When reared on *L. grayi*, *P. zelacon* pupate at lower masses than *P. oregonius* larvae. The few *P. zelacon* larvae that survive to pupation on *A. dracunculus* pupate at lower masses than *P. oregonius* larvae. Pupal mass of hybrids is significantly higher for larvae reared on *L. grayi* than on *A. dracunculus*, in females than in males, and in OZ crosses than in ZO crosses in both series 1 and 2 (all $P \leq 0.05$, ANOVA). There is a significant interaction effect between host species and sex of the larva ($P \leq 0.05$ for both series) (Fig. 3). The differences between reciprocal crosses in both females and males indicate that non-genetic maternal effects may affect pupal masses in hybrid larvae developing on these two hosts.

Overall, these three measures of larval performance suggest that hybrids receive alleles from both parental species in a way that allows male and female larvae to survive on both parental hosts. The results provide no indication that the genes for larval performance are on the X chromosome, as are the genes for host preference in ovipositing females of these species. Consequently, neither pleiotropy nor physical linkage of loci for oviposition preference and larval performance seem to be involved in any major way in the evolution of interspecific differences in host use between these two species. Moreover, the results indicate that, unlike oviposition preference, Y-linked or non-genetic maternal effects may influence one or more components of larval performance on different hosts. These maternal effects cannot be due to differences in the diets of the parents, because all parents were reared on the same plant species. Additional information on the source of these maternal effects could be obtained from F2 and backcross progeny, but viable offspring have been very difficult to obtain from such crosses using these populations²⁸. Finally, the results show that the interpretation of the genetics of larval performance varies with the specific measure of performance, the host plant species used in comparing reciprocal crosses, and the sex of the larvae. □

Received 30 October 1989; accepted 18 January 1990.

- Janzen, D. H. *Oxford Surveys in Evolutionary Biology* **1**, 85–140 (1984).
- Strong, D. R., Lawton, J. H. & Southwood, R. *Insects on Plants: Community Patterns and Mechanisms* (Blackwell, Oxford, 1984).
- Ehrlich, P. R. & Raven, P. H. *Evolution* **18**, 586–608 (1964).
- Southwood, T. R. E. in *Insect/Plant Relationships* (ed. van Emden, H. F.) 3–30 (Blackwell, Oxford, 1973).
- Lawton, J. H. & Schroeder, D. *Nature* **265**, 137–140 (1977).
- Price, P. W. *Evolutionary Biology of Parasites* (Princeton University Press, 1980).
- Mitter, C., Farrell, B. & Wiegmann, B. *Am. Nat.* **231**, 107–128 (1988).
- Bush, G. L. in *Evolutionary Strategies of Parasitic Insects and Mites* (ed. Price, P. W.) 187–206 (Plenum, London, 1975).
- Felsenstein, J. *Evolution* **35**, 124–138 (1981).
- Rauscher, M. D. *Evolution* **38**, 596–608 (1984).
- Diehl, S. R. & Bush, G. L. in *Speciation and its Consequences* (eds Otte, D. & Endler, J.) 345–365 (Sinauer Associates, Sunderland, Massachusetts, 1989).
- Via, S. *Evolution* **40**, 778–785 (1986).
- Jaenike, J. *Heredity* **59**, 363–369 (1987).
- Lande, R. *Genetics* **94**, 203–215 (1980).
- Futuyma, D. J. & Peterson, S. C. *A. Rev. Entomol.* **30**, 217–238 (1985).
- Bush, G. L. in *Evolutionary Behavior Genetics* (ed. Huettel, M. D.) 1–5 (Plenum, New York, 1987).
- Butlin, R. *Trends Ecol. Evol.* **2**, 309–311 (1987).
- Futuyma, D. J. in *Evolutionary Behavior Genetics* (ed. Huettel, M. D.) 295–302 (Plenum, New York, 1987).
- Courtney, S. P. & Chen, G. K. *Funct. Ecol.* **2**, 521–528 (1988).
- Futuyma, D. J. & Moreno, G. A. *Rev. Ecol. Syst.* **19**, 207–233 (1988).
- Ng, D. *Nature* **334**, 611–612 (1988).
- Singer, M. C., Ng, D. & Thomas, C. D. *Evolution* **42**, 977–985 (1988).
- Thompson, J. N. *Entomol. exp. Appl.* **47**, 3–14 (1988).
- Tauber, C. A. & Tauber, M. J. *Evol. Ecol.* **1**, 175–186 (1987).
- Courtney Smith, D. *Nature* **336**, 66–67 (1988).
- Feder, J. L., Chalcote, C. A. & Bush, G. L. *Nature* **336**, 61–64 (1988).
- McPherson, B. A., Courtney Smith, D. & Berlocher, S. H. *Nature* **336**, 64–66 (1988).
- Thompson, J. N. *Evolution* **42**, 1223–1234 (1988).
- Miller, J. S. *Bull. Am. Mus. nat. Hist.* **186**, 365–512 (1987).
- Scriber, J. M. *Tokurana (Acta Rhopal.)* **6/7**, 1–50 (1984).
- Wiklund, C. *Oikos* **36**, 163–170 (1981).
- Berenbaum, M. *Evolution* **37**, 163–179 (1983).
- Clarke, C. A. & Larsen, T. B. *Syst. Entomol.* **11**, 175–181 (1986).
- Thompson, J. N. *Evolution* **42**, 118–128 (1988).
- Feeny, P. in *Herbivory: Tropical and Temperate Perspectives* (eds Price, P. W., Lewinsohn, T. M., Benson, W. W. & Fernandes, G. W.) (Wiley, New York, in the press).
- Scriber, J. M., Lederhouse, R. C. & Hagen, R. H. in *Herbivory: Tropical and Temperate Perspectives* (eds Price, P. W., Lewinsohn, T. M., Benson, W. W. & Fernandes, G. W.) (Wiley, New York, in the press).
- White, M. J. D. *Animal Cytology and Evolution* (Cambridge University Press, 1973).

ACKNOWLEDGEMENTS. We thank Hal Hansel and Kirsten Price for help with the experiments, and Hal Hansel, Amitabh Joshi, Lisa Valburg and Sara Via for comments on the manuscript. This work was supported by the NSF and the USDA Competitive Grants Program.

Host-parasitoid associations in patchy environments

S. W. Pacala*, M. P. Hassell† & R. M. May†

* Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut 06268, USA
† Department of Biology and Centre for Population Biology, Imperial College, Silwood Park, Ascot, Berks SL5 7PY, UK

STUDIES of insect host-parasitoid interactions have contributed much to the consensus that spatial patchiness is important in the regulation of natural populations^{1–5}. A variety of theoretical models predict that host and parasitoid populations, although unstable in the absence of environmental heterogeneity, may persist at roughly steady overall densities in a patchy environment owing to variation in levels of parasitism from patch to patch. Observed patterns of parasitism, however, have a variety of forms (with variation in attack rates among patches depending directly or indirectly on host density, or showing variation uncorrelated with host density). There is some confusion about the dynamical consequences of these different forms^{6,7}. Here we first show how the dynamical effects of all these forms of environmental heterogeneity can be assessed by a common criterion. This 'CV² > 1 rule' states that the overall population densities will remain roughly steady from generation to generation if the coefficient of variation squared (CV²) of the density of searching parasitoids in the vicinity of each host exceeds approximately unity. By partitioning CV² into components, we show that both direct and inverse patterns of dependence on host density, and density-independent patterns, all contribute to population regulation in the same way. Second, we show how a maximum-likelihood method can be applied to the

kind of field data that are usually available (that is, percentage parasitism versus local host density) to estimate the components of CV^2 . This analysis indicates that heterogeneity is large enough to stabilize dynamics in 9 of 34 published studies, and that density-independent heterogeneity is the main factor in most cases.

Insect parasitoids lay their eggs on, in, or close to, the bodies of other arthropods (usually insects), which are eventually killed by the feeding parasitoid larvae. The many laboratory and field studies showing levels of percentage parasitism by insect parasitoids in relation to local host density per 'patch' are reviewed in refs 8–10, and categorized as 'direct' (Fig. 1a), 'inverse density-dependent' (Fig. 1b) or 'density-independent' (Fig. 1c, d). It has been assumed that, of these patterns, direct density-dependent patterns are the most important in promoting population regulation. More recently, however, it has been emphasized that inverse patterns^{10,11}, and even density-independent ones^{2,3}, can also be important contributors to population regulation. As a result, the dynamical consequences of observed spatial patterns of parasitism cannot be simply inferred from the shape of the relationships. Here we present a criterion governing the overall dynamics of associations in which the interactions between hosts and parasitoids are of predominant importance (that is, where other density-dependent effects on host or parasitoid fecundity, or both, or survival are relatively unimportant or absent). The criterion is an approximate one, but is simple, general, and can be readily evaluated by using the kinds of data that are usually collected in field studies.

We begin with the widely studied general host-parasitoid model¹²

$$N_{t+1} = \lambda N_t F(N_t, P_t) \quad (1a)$$

$$P_{t+1} = q N_t [1 - F(N_t, P_t)] \quad (1b)$$

where N_t and P_t are, respectively, the host and parasitoid population sizes in generation t , λ is the finite rate of increase of the host (which is taken to be independent of host density; stability considerations will obviously be more complex in situations where λ is significantly density-dependent), $F(N_t, P_t)$ is the fraction of hosts that escape parasitism, and q is the average number of female parasitoids emerging from each host parasitized. If the environment is patchy, $F(N_t, P_t)$ is calculated as the average—over all hosts in all patches—of the probability of escaping parasitism. This probability is taken to be $\exp(-ap)$, where a is the *per capita* searching efficiency of the parasitoids and p is the local (within patch) population density of searching parasitoids. The local abundance of parasitoids may vary from place to place in the habitat, and the mean of p may be a function of mean host density, mean parasitoid density and local host density. Note that $\exp(-ap)$ is the probability of

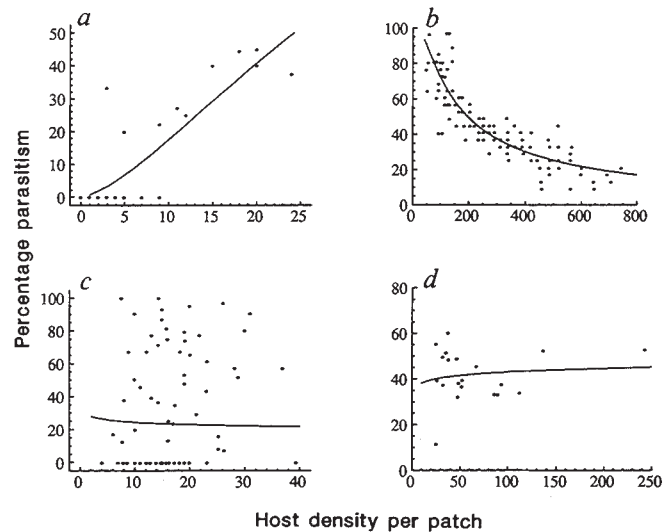


FIG. 1. Examples of field studies showing percentage parasitism as a function of host density. Curves are means predicted by the model when evaluated at maximum-likelihood estimates. *a*, Direct density dependence from ref. 29. Because the calculated $CV^2 > 1$ ($CV^2 = 1.52$; see Table 1), the models predict there is sufficient heterogeneity to stabilize the interacting populations. Furthermore, because C_D is significantly > 2 , μ is significantly positive and C_I is close to 1 (see estimates and confidence limits in Table 1), virtually all of the stabilizing heterogeneity in this example results from the aggregation of parasitoids on patches of high host density. *b*, Inverse density dependence from ref. 23. Unlike the example in *a*, heterogeneity is now not sufficiently large to stabilize the dynamics ($CV^2 = 0.37$, which is < 1). *c*, Illustration of a case (ref. 25) with considerable density-independent heterogeneity, but virtually no density-dependent heterogeneity. Because C_I is significantly > 2 (shown by the lower confidence bound on σ^2), CV^2 is significantly > 1 . Therefore the density-independent heterogeneity in *c* is apparently large enough by itself to stabilize the dynamics of the model in equation (1). Although the data in *c* may seem erratic, they actually contain more evidence of factors that can regulate the host and parasitoid populations than the previous examples. *d*, Illustration of a case (ref. 33) with virtually no stabilizing heterogeneity. Because μ is near zero ($\mu = 0.07$), there is little density-dependent heterogeneity, and because σ^2 is small, there is also little of the 'greater-than-binomial' variance that leads to the prediction of stability in *c*.

escaping attack from parasitoids that search independently and randomly within a patch.

Hassell and May² suggested that the coefficient of variation squared of p (that is, $CV^2 = (\text{variance}/\text{mean}^2)$) could provide a guide to the dynamical behaviour of equations (1a) and (1b), and showed for a particular example that the host and parasitoid populations were stable if $CV^2 > 1$. We have now demonstrated

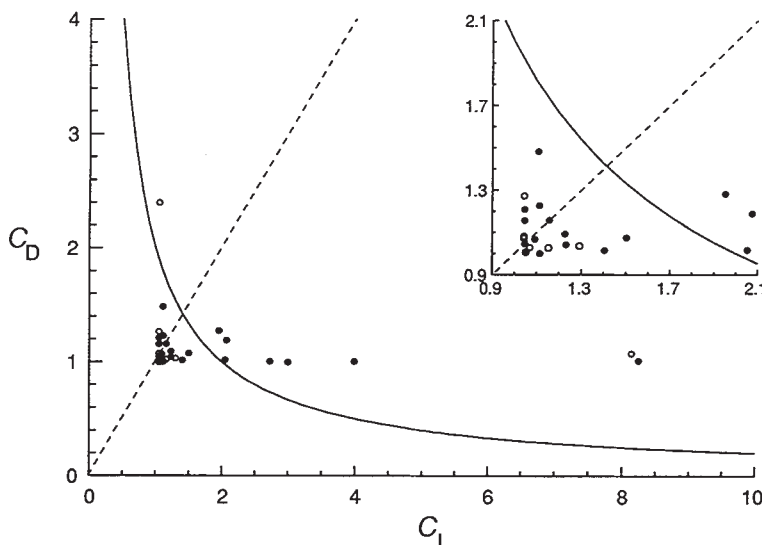


FIG. 2. Host density-independent heterogeneity (C_I) plotted against host density-dependent heterogeneity (C_D). Each solid point represents a different pair of species, whereas each hollow point comes from the same pair of species as one of the solid points. The solid curve is the contour where $CV^2 = 1$. Dynamics are predicted to be stable if a point lies above this line (thus $CV^2 > 1$). The dashed line separates points for which $C_D > C_I$ from those for which $C_D < C_I$. The inset magnifies the cluster of points nearest the origin.

TABLE 1 Parameter estimates and components of heterogeneity

Ref.	$\sigma^2 \pm 95\% \text{ c.l.}^\dagger$	$\mu \pm 95\% \text{ c.l.}$	C_D (=HDD)	C_I (=HDI)	CV^2	Host $CV^2(V^2)$	Host species	Parasitoid species
23	0.113 ± 0.050	-0.953 ± 0.086	1.228	1.113	0.367	0.251	<i>Lymantria</i> (=Porthetria) <i>dispar</i>	<i>Ooencyrtus kuwanai</i>
24	1.074 ± 0.450	-0.769 ± 0.166	1.190	2.074	1.468	0.322	<i>Callosobruchus chinensis</i>	<i>Anisopterimalus calandreae</i>
24	0.956 ± 0.245	-0.793 ± 0.135	1.274	1.956	1.492	0.436	<i>C. chinensis</i>	<i>Heterospilus prosopidis</i>
25	2.994 ± 1.417	0.247 ± 0.968	1.006	3.994	3.019	0.101	<i>Rhopalomyia californica</i>	<i>Torymus baccaridis</i>
25	7.246 ± 3.271	-0.238 ± 0.968	1.011	8.246	7.333	0.197	<i>R. californica</i>	<i>Tetrastichus</i> sp.
26	0.050 ± 0.068	-0.200 ± 0.654	1.007	1.050	0.057	0.175	<i>Silo pallipes</i>	<i>Agriotypus armatus</i>
27	0.233 ± 0.127	0.200 ± 0.234	1.041	1.233	0.284	1.034	<i>Sceliphron assimile</i>	<i>Mellitobia chalybii</i>
28	0.230 ± 0.167	0.574 ± 0.453	1.092	1.230	0.343	0.279	<i>Phytomyza ilicis</i>	<i>Chrysocharis gemma</i>
29	0.109 ± 0.170	0.913 ± 0.311	1.483	1.109	0.645	0.579	<i>Delia radicum</i>	<i>Trybliographa rapae</i>
29*	0.050 ± 0.084	1.460 ± 0.255	2.401	1.050	1.521	0.657	<i>D. radicum</i>	<i>T. rapae</i>
29*	0.050 ± 0.047	0.669 ± 0.228	1.267	1.050	0.330	0.597	<i>D. radicum</i>	<i>T. rapae</i>
30	0.050 ± 0.975	-0.689 ± 0.824	1.156	1.050	0.214	0.329	<i>Papilio xuthus</i>	<i>Trichogramma papilionis</i>
‡	0.115 ± 0.453	-0.047 ± 0.513	1.001	1.115	0.116	0.359	<i>Chirosia histicrina</i>	Total parasitism
31	0.160 ± 0.148	0.874 ± 0.302	1.159	1.160	0.344	0.209	<i>Fiorinia externa</i>	<i>Aspidiotiphagus citrinus</i>
31*	0.077 ± 0.044	0.395 ± 0.085	1.024	1.077	0.103	0.153	<i>F. externa</i>	<i>A. citrinus</i>
31*	0.050 ± 0.070	0.625 ± 0.175	1.077	1.050	0.131	0.197	<i>F. externa</i>	<i>A. citrinus</i>
32	0.504 ± 0.316	1.030 ± 0.950	1.076	1.504	0.608	0.071	<i>Epinotia tedella</i>	<i>Apanteles tedellae</i>
32	0.404 ± 0.047	-0.434 ± 0.887	1.016	1.404	0.424	0.073	<i>E. tedellae</i>	<i>Pimpla dubius</i>
33*	0.050 ± 0.071	0.363 ± 0.275	1.068	1.050	0.121	0.513	<i>Parlatoria oleae</i>	<i>Aphytis paramaculicornis</i>
33	0.050 ± 0.051	0.070 ± 0.210	1.003	1.050	0.053	0.613	<i>P. oleae</i>	<i>Coccophagoides utilis</i>
33	0.050 ± 0.056	0.513 ± 0.336	1.047	1.050	0.099	0.177	<i>P. oleae</i>	<i>A. paramaculicornis</i>
34*	0.162 ± 0.125	0.159 ± 0.185	1.022	1.161	0.187	0.865	<i>Aonidiella aurantii</i>	<i>Aphytis melinus</i>
34	2.000 ± 0.622	-0.025 ± 0.577	1.000	3.000	2.000	0.475	<i>A. aurantii</i>	<i>A. melinus</i>
34*	0.050 ± 0.065	0.235 ± 0.659	1.006	1.050	0.056	0.104	<i>A. aurantii</i>	<i>A. melinus</i>
34*	0.060 ± 0.162	-0.036 ± 0.751	1.000	1.060	0.060	0.126	<i>A. aurantii</i>	<i>A. melinus</i>
35	0.050 ± 0.069	0.931 ± 0.554	1.212	1.050	0.273	0.244	<i>Eupteryx cyclops</i>	<i>Anagrus</i> sp.
							<i>E. urticae</i>	
36	1.724 ± 1.316	0.079 ± 0.074	1.003	2.724	1.732	0.461	<i>Polistes exclamans</i>	<i>Elasmus polistis</i>
36*	7.143 ± 3.832	-0.463 ± 0.269	1.077	8.143	7.770	0.357	<i>P. exclamans</i>	<i>E. polistis</i>
37	0.050 ± 0.065	0.008 ± 0.230	1.000	1.050	0.050	0.312	<i>Ichneumon purchasi</i>	<i>Cryptochaetum iceryae</i>
37	0.050 ± 0.075	0.047 ± 0.284	1.011	1.050	0.062	0.566	<i>I. purchasi</i>	<i>C. iceryae</i>
37*	0.299 ± 0.350	-0.254 ± 0.503	1.031	1.299	0.339	0.480	<i>I. purchasi</i>	<i>C. iceryae</i>
37*	0.050 ± 0.071	-0.014 ± 0.151	1.000	1.050	0.050	0.789	<i>I. purchasi</i>	<i>C. iceryae</i>
38	1.053 ± 3.223	0.175 ± 0.660	1.017	2.052	1.087	0.560	<i>Trypargilum politum</i>	<i>Mellitobia</i> sp.
39	0.093 ± 0.027	-0.531 ± 0.124	1.068	1.093	0.167	0.240	<i>Lymantria</i> (=Porthetria) <i>dispar</i>	<i>Ooencyrtus kuwanai</i>

† All estimates of 0.05 for σ^2 should actually lie between 0 and 0.05 because this is the lower bound of our algorithm's resolution. Note that this biases corresponding values of C_I and $CV^2 > 1$ by $< 5\%$. Asterisks indicate replicate data sets; see text.

‡ J. H. Lawton, unpublished data.

that this ' $CV^2 > 1$ rule' is indeed an approximate criterion for stability in a variety of models, some general and some specific, provided that CV^2 is calculated with respect to a randomly chosen host (not host patch). Choosing two explicit examples from those discussed in detail elsewhere¹³, we can show that: (1) the $CV^2 > 1$ rule is exact for all λ if adult parasitoid density varies from patch to patch as a gamma-distributed variable; and (2) the rule is a good approximation if the mean value of p is $P_f(N)$ where N is the local density of hosts, assuming any functional form $f(N)$ and for any probability-density functions for N and p .

Here we show how the $CV^2 > 1$ rule can be applied by using information on the actual distribution of parasitoid attacks, as observed in the field. We begin by making the reasonable assumption that the local parasitoid density can be described by a regression function: $p = P_f(N)\varepsilon$, where ε is a random variable with unit mean and denotes the residual density-independent variation. By expanding the density-dependent function $f(N)$ in Taylor series to first-order about the mean at $N = N_i$, $CV^2 > 1$ can be approximated as:

$$CV^2 \approx C_I C_D - 1; \quad C_I = 1 + \sigma^2; \\ C_D = 1 + V^2 \left[\frac{df(N_i)/dn}{f(N_i)} \right]^2 / f(N_i)^2 \quad (2)$$

Here σ^2 is the variance of ε , n is the host density relative to the local mean density ($n = N/N_i$), and V is the coefficient of variation of the local host density calculated with respect to a randomly chosen host. Specifically, if $p(N)$ is the fraction of patches containing N hosts, then $p(N)N/N_i$ is the fraction of hosts that inhabit a patch containing N hosts. V is calculated using $p(N)N/N_i$. C_I represents the component of heterogeneity that is independent of host density, and C_D represents the component that depends on host density.

Note from equation (2) that C_I and C_D both affect $CV^2 > 1$ in the same way. Also note that because the slope-determining expression $df(N_i)/dn$ in C_D is squared, the effects of direct (positive slope) and inverse (negative slope) dependence on local host density are identical. Therefore all of the patterns illustrated in Fig. 1 contribute to population regulation in the same way, and to a degree that depends on the magnitude of the CV^2 of the searching parasitoids.

Although there are few data on the distribution of searching parasitoids in the field (but see refs 14–17), there are many studies reporting percentage parasitism as a function of local host density (see Fig. 1 and Table 1). We have developed a maximum-likelihood method that allows us to estimate $df(N_i)/dn$ and σ^2 from these widely available data. After estimating V in the usual way from the data on local host densities, we can then calculate CV^2 , C_I and C_D using equation (2). The likelihood estimator assumes that ε is gamma distributed and that $f(N) = c(N/N_i)^\mu$ (so that $[df(N_i)/dn]/f(N_i) = \mu$). The assumption of a gamma distribution is reasonable in light of the few studies that have measured the distribution of attacks per host^{18–20}. We have chosen the function $c(N/N_i)^\mu$ for $f(N)$ because this has been widely used in the theoretical literature^{11,21} and provides better fits than does a linear function. Details of the estimation procedure are given elsewhere²².

The data on percentage parasitism versus local host density per patch come largely from two recent reviews^{9,10}. Of 86 data sets initially screened, we selected 34 that provide sufficient information to use our estimation procedure. For each of these 34 sets, we obtained estimates of σ^2 , μ and V , and then used equation (2) to calculate C_I , C_D and CV^2 . These estimates are listed in Table 1. Some of the 34 data sets are temporal or spatial replicates of others. We therefore also produced a reduced assemblage of 22 data sets by choosing a single replicate at

random and omitting all others; the 12 omitted sets are designated by an asterisk in the first column in Table 1.

The curves in Fig. 1 give the mean percentage parasitism predicted by the model when evaluated at the maximum-likelihood estimates. Note the close correspondence between the trends in the data and the fitted means. Figure 2 is a plot of C_D versus C_I for all 34 data sets, with the $CV^2 = 1$ contour overlaid. The figure shows that heterogeneity is large enough for a prediction of stability ($CV^2 > 1$) in 9 of the 34 cases (assuming, of course, that these levels of heterogeneity are repeated from generation to generation). It is of interest that $C_I > C_D$ in eight of these nine cases, and that density-independent heterogeneity is large enough to stabilize dynamics by itself ($C_I > 2$) in seven of the nine (statistically significant in four of these seven; shown by the lower 95% confidence limit of σ^2). The result—that host density-independent heterogeneity is large enough to stabilize dynamics more often than is host density-dependent heterogeneity—is not due to any pseudo-replication in our 34 data sets, because it is also obtained from the assemblage of 22 data sets, each representing a different pair of species (solid points in Fig. 2). Notice that $CV^2 > 1$ in seven of the 22 sets, $C_I > C_D$ in seven of these seven, and density-independent heterogeneity is large enough to stabilize dynamics ($C_I > 2$) in six of seven.

In many of the cases in which $CV^2 < 1$, heterogeneity is still large enough to contribute substantially to population regulation. The data in Fig. 2 show that density-dependent and -independent heterogeneity are about equally important in those examples in which $CV^2 < 1$. Also, the estimates of μ in Table 1 show that, where density-dependent heterogeneity is important, it is caused by inverse density dependence as often as it is caused by direct density dependence.¹¹

Environmental heterogeneity has often been regarded as a complicating factor in population dynamics, and one that rapidly leads to analytical intractability. The study presented here indicates that this need not be so. The $CV^2 > 1$ rule explains the consequences of heterogeneity for population dynamics in terms of a simple description of the heterogeneity itself. The rule gives a rough prediction of the effects of heterogeneity and also identifies the kinds of heterogeneity that contribute to population regulation. It also offers the hope that comparable criteria and methods can be found for assessing the role of environmental heterogeneity in a much broader range of interactions among species. □

Received 10 July 1989; accepted 10 January 1990.

- Hassell, M. P. & May, R. M. *J. Anim. Ecol.* **43**, 567–594 (1974).
- Hassell, M. P. & May, R. M. *Ann. zool. Fenn.* **25**, 55–61 (1988).
- Chesson, P. L. & Murdoch, W. W. *Am. Nat.* **127**, 696–715 (1986).
- Reeve, J. D. *Am. Nat.* **132**, 810–836 (1988).
- Taylor, A. D. *Am. Nat.* **132**, 417–436 (1988).
- Ehler, L. E. *Environ. Ent.* **15**, 1268–1271 (1986).
- Hassell, M. P. *Ecol. Ent.* **7**, 365–377 (1982).
- Lessells, C. M. *J. Anim. Ecol.* **54**, 27–41 (1985).
- Stiling, P. D. *Ecology* **68**, 844–856 (1987).
- Walde, S. J. & Murdoch, W. W. *A. Rev. Ent.* **33**, 441–466 (1988).
- Hassell, M. P. *IMA J. appl. med. Biol.* **1**, 33 (1984).
- Hassell, M. P. *The Dynamics of Arthropod Predator–Prey Systems* (Princeton University Press, 1978).
- Hassell, M. P., May, R. M., Pacala, S. & Chesson, P. *Am. Nat.* (in the press).
- Waage, J. K. *Ecol. Ent.* **8**, 447–453 (1983).
- Smith, A. D. M. & Maelzer, D. A. *Ecol. Ent.* **11**, 425–434 (1986).
- Casas, J. *Physiol. Ent.* **13**, 373–380 (1988).
- Casas, J. *Ecol. Ent.* **14**, (1989).
- Griffiths, K. J. *Can. Ent.* **101**, 673–713 (1969).
- Griffiths, K. J. & Holling, C. S. *Can. Ent.* **101**, 785–818 (1969).
- May, R. M. *J. Anim. Ecol.* **47**, 833–843 (1978).
- Hassell, M. P. & May, R. M. *J. Anim. Ecol.* **42**, 693–726 (1973).
- Pacala, S. & Hassell, M. P. *Am. Nat.* (in the press).
- Brown, M. W. & Cameron, E. A. *Environ. Ent.* **8**, 77–80 (1979).
- Hassell, M. P., Lessells, C. M. & McGavin, G. C. *Ecol. Ent.* **10**, 393–402 (1985).
- Ehler, L. E. *Environ. Ent.* **15**, 1268–1271 (1986).
- Elliott, J. M. *J. Anim. Ecol.* **52**, 315–330 (1983).
- Freeman, B. E. & Parnell, J. R. *J. Anim. Ecol.* **42**, 779–784 (1983).
- Heads, P. A. & Lawton, J. H. *Oikos* **40**, 267–276 (1983).
- Jones, T. H. & Hassell, M. P. *Ecol. Ent.* **13**, 309–317 (1988).
- Hirose, Y., Kimoto, H. & Hiemata, K. *Appl. Ent. Zool.* **11**, 116–125 (1976).
- McClure, M. S. *Environ. Ent.* **6**, 551–555 (1977).
- Munster-Svensden, M. *Ecol. Ent.* **5**, 373–383 (1980).

- Murdoch, W. W., Reeve, J. D., Huffaker, C. E. & Kennett, C. E. *Am. Nat.* **123**, 92 (1984).
- Smith, A. D. M. & Maelzer, D. A. *Ecol. Ent.* **11**, 425–434 (1986).
- Stiling, P. D. *J. Anim. Ecol.* **49**, 793–805 (1980).
- Strassman, J. E. *Ecology* **62**, 1225–1233 (1981).
- Thorarinsson, K. *Ecology* (in the press).
- Trexler, J. C. *Oikos* **44**, 415–422 (1985).
- Weseloh, R. M. *Ann. ent. Soc. Am.* **65**, 64–69 (1972).

ACKNOWLEDGEMENTS. We thank Hefin Jones for help in compiling the data, and Charles Godfray and John Reeve for comments on the manuscript. This work was supported by the NERC Centre for Population Biology at Imperial College, the NERC (M.P.H.) and The Royal Society (R.M.M.).

Transparency and coherence in human motion perception

G. R. Stoner, T. D. Albright & V. S. Ramachandran*†

The Salk Institute for Biological Studies, La Jolla, California 92037, USA
* Department of Psychology, C-009 University of California at San Diego, La Jolla, California 92093, USA

WHEN confronted with moving images, the visual system often must decide whether the motion signals arise from a single object or from multiple objects^{1–5}. A special case of this problem arises when two independently moving gratings are superimposed. The gratings tend to cohere and move unambiguously in a single direction² (pattern motion) instead of moving independently (component motion). Here we report that the tendency to see pattern motion depends very strongly on the luminance of the intersections (that is, to regions where the gratings overlap) relative to that of the gratings in a way that closely parallels the physics of transparency. When the luminance of these regions is chosen appropriately, pattern motion is destroyed and replaced by the appearance of two transparent gratings moving independently. The observations imply that motion detecting mechanisms in the visual system must have access to tacit 'knowledge' of the physics of transparency and that this knowledge can be used to segment the scene into different objects. The same knowledge could, in principle, be used to avoid confusing shadows with real object boundaries.

A simple example of pattern motion can be seen in the complex plaid pattern shown in Fig. 1c. This pattern was created by superimposing two identical square wave gratings (Fig. 1a and b). The motion of each component grating is indicated by the arrows and this is what human observers always see when either grating is viewed separately. When viewing the plaid pattern, however, observers usually report seeing upward pattern motion that is (Fig. 1c) different from either of the two component directions of motion. Perhaps the visual system computes the loci of possible motion for the two gratings separately and then determines the single point where the loci intersect². This point would uniquely specify both the direction and velocity of the coherently moving plaid pattern.

To explore the role of transparency in motion perception, we created a new class of moving stimuli that convey a striking impression of perceptual transparency (Fig. 2b). We found that this was best achieved by using asymmetrical square wave gratings instead of sine waves (see the legend of Fig. 2a for details). Although sine-wave gratings have been used in previous studies of motion coherence^{2,5,6}, we found that identical effects are obtainable with square-wave gratings. Also, it is far easier to manipulate luminance levels in square-wave gratings to achieve transparency effects. Asymmetric gratings (duty cycle is narrow bar width/(narrow bar + wide bar) and equals 0.286) were used to bias the figure-ground interpretation, which has, in turn, a strong effect on the interpretation of transparency. In these patterns a specific cycle portion (namely the narrow bars) was consistently seen as foreground. Trials consisted of brief

† To whom correspondence should be addressed.

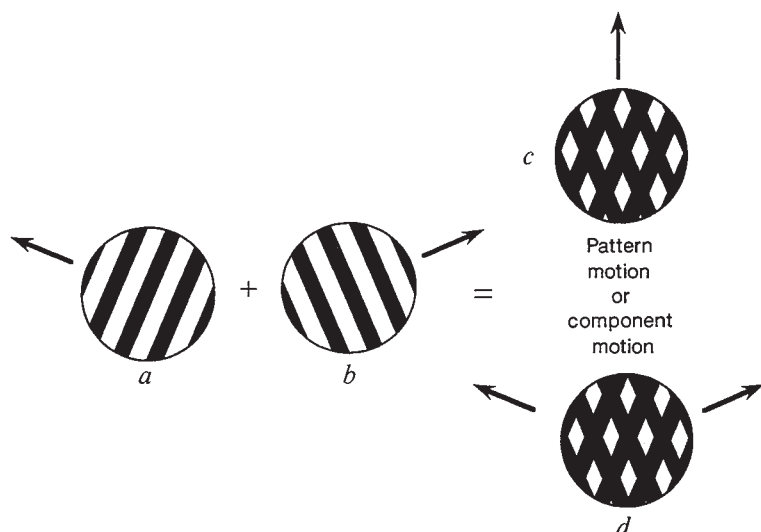


FIG. 1 *a* and *b*, Two square wave gratings. Each grating is moved orthogonally to its orientation and the velocities of the two gratings are identical. The perceived directions of motion are indicated by the arrows. *c*, Superimposition of *a* and *b*. Instead of seeing the two gratings move in separate directions, observers usually see a 'plaid' moving in a single direction (as shown by the arrow). *d*, Under some conditions, the gratings are seen to slide past each other, that is, component motion is seen instead of pattern motion.

presentations (1.5 s), after which the subject had to indicate (with a key press) whether he saw pattern motion. Subjects were instructed to fixate on a small cross in the centre of the aperture for the duration of each trial. Subjects under this type of instruction are capable of stable and reliable fixation⁷.

On each trial the display consisted of three regions: (1) that formed by the intersection of the two gratings; (2) the narrow bars of both gratings (which were identical and held constant at 90 cd m^{-2}); and (3) the wide bars (which we term background) of both gratings (which were held constant at 231 cd m^{-2}). Perceptual transparency was manipulated by varying only the luminance of the intersections. Fifteen intersection luminance values were presented on a pseudo-random schedule. These varied from 4.90 to 125.3 cd m^{-2} in increments of $\sim 8.50 \text{ cd m}^{-2}$.

The rules of perceptual transparency have been previously studied for stationary objects⁸⁻¹⁰. To understand these rules imagine two gratings superimposed on a background whose luminance is 100 cd m^{-2} . If the gratings are neutral density filters

of 50% transmittance, then their luminances would be 50 cd m^{-2} , but the luminance of the intersections would be 50% of 50 cd m^{-2} , or 25 cd m^{-2} , so the relationship would be multiplicative rather than additive. Any intersection luminance above 25 cd m^{-2} but below 50 cd m^{-2} would still be compatible with the physics of transparency but the gratings would then be seen also to have some surface reflectance of their own, that is, they would both look like translucent or frosted glass plates. Values below 25 and above 50 cd m^{-2} , on the other hand, would be incompatible with physics. In our moving displays, these rules of transparency dictate that the optimal conditions for pure transparency (equivalent to the neutral density filter case) could be calculated by multiplying the narrow bar-to-background ratio (constant: 0.390) by the narrow bar luminance (90 cd m^{-2}): this would be 35 cd m^{-2} . Perceptual transparency should also be seen when the intersection luminance is increased beyond this point but remains less than the narrow bar luminance (Fig. 2*b*). Within this transparency zone, the foreground grating should

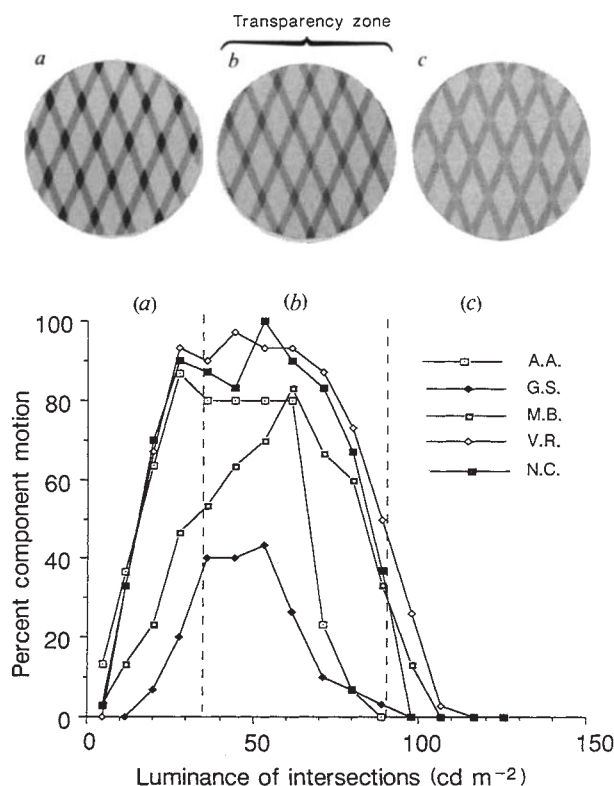


FIG. 2 *a*, Example of stimuli used in our experiment. The intersection luminance in this stimulus is too dark to be compatible with the physics of transparency. This leads to a decline in component motion (Fig. 3). Our stimuli were generated using a high-resolution graphics display controller (Pepper SGT, Number Nine Computer Corporation: 640×480 pixels, 8 bits per pixel; 60 Hz, non-interlaced) operating in an AT computer. Stimuli were displayed on a 14-inch analog RGB video monitor (Zenith ZCM-1490, flat technology CRT) and were viewed through a circular aperture subtending 11° at a distance of 57 cm. They were moved by updating their position in synchrony with the vertical refresh of the monitor on alternate cycles (that is, every 33.3 ms). *b*, Similar to *a*, except that the intersections between the two gratings are only slightly darker than other regions in the image and this conveys a vivid impression of transparency. When this pattern is moved, subjects usually see the two gratings sliding past each other, that is, they see component motion rather than coherent pattern motion in a single direction. *c*, Identical to *a*, except that the luminance of the intersections is brighter than that of either grating. This stimulus is physically incompatible with two transparent gratings and is usually seen as a single coherently moving plaid. It lies to the right of the 'transparency zone' in Fig. 3.

FIG. 3 Probability of component motion percept as a function of plaid pattern intersection luminance. Both gratings were of the same spatial frequency (1.75 cycles per deg). On each trial the individual gratings were moved at an angle of 135° relative to one another at a speed of 3° s^{-1} , resulting in a pattern speed of 8° s^{-1} . Pattern direction was either up or down, and varied on a random schedule. Each datum point represents the mean of 30 trials at each intersection value (in three series of 10 trials per intersection value). Data are shown from three naive (A.A., M.B., N.C.) and two experienced subjects (G.S., V.R.). The intersection/luminance was varied in roughly equal steps. The 'transparency zone' extends from pure transparency (35 cd m^{-2}) up to the point of 'occlusion' (90 cd m^{-2}). Component motion is most likely within a region roughly centred on the transparency zone.

be perceived to be transparent but also having some surface reflectance. When both gratings are of the same luminance (as in these experiments), an intersection luminance equal to the narrow bar luminance is the only condition compatible with occlusion. It follows that intersection luminances of $<35 \text{ cd m}^{-2}$ (Fig. 2a) or $>90 \text{ cd m}^{-2}$ (Fig. 2c) are totally incompatible with physical transparency or occlusion. They are unlikely to lead an observer to interpret the display as two independently moving gratings and we therefore predict a decline in component motion judgements under these conditions.

Figure 3 shows the results: notice that for each subject there is a range of intersection luminances for which there is a high probability of seeing component motion. The midpoint of this range coincides with the centre of the transparency zone (the region in which the gratings both filter and reflect light). This implies that the tendency to see component motion depends strongly on whether or not the gratings look transparent. If the luminance of the intersections is adjusted so that the gratings look transparent, component motion is usually seen instead of pattern motion. On the other hand, if the luminance ratios are incompatible with two physically transparent gratings, subjects are much less likely to see component motion. For example, Fig. 2c is physically incompatible with transparency; there are no two gratings that can overlap to produce intersections that are actually brighter than either grating alone and consequently, the visual system rejects this percept.

Another interpretation of our findings might be that adding luminance to the intersections would introduce new horizontally oriented Fourier components, whose unambiguous upward motion might capture the two gratings^{4,5} and result in pattern motion. This model would predict that maximal component motion should occur when the two gratings are simply added together, but our findings contradict this. In fact, our data (Fig. 3) demonstrate that component motion is much greater when the interaction between the grating luminances is multiplicative (as in two overlapping neutral density filters; corresponding to the left-most dotted line in Fig. 3) rather than additive. In Fig. 3 the intersection luminance value closest to the linear (additive) superimposition is 4.90 cd m^{-2} (leftmost data points). Horizontally oriented Fourier components are at a minimum here and increase in power as the intersection luminance increases. Thus, we might expect component motion to decrease monotonically as we move towards the right in the graph, but instead it increases until the value becomes incompatible with transparency and then declines again. These findings are more consistent with a transparency model than with an interpretation in terms of Fourier components.

Our results have two interesting implications. First, if computational models of motion perception are to provide more than "mere caricatures"¹¹ of human visual processes, they have to take into account the effects of multiple sources of information^{12,13}, such as those implied by the transparency illusion reported here. Second, a certain proportion of cells in the middle temporal area of primates¹⁴⁻¹⁶ exhibit direction-selective responses to pattern motion¹⁷⁻¹⁹, rather than component motion, when confronted with Fig. 1c. Would this response to pattern motion be reduced if the gratings were made to look transparent? □

13. Ramachandran, V. S. in *AI and the Eye* (ed. Blake, A.S. & Troscianko, T.) (Wiley, Chichester, 1990).
14. Zeki, S. M. *Proc. R. Soc. B* **207**, 239-248 (1980).
15. Allman, J., Miezin, F. & McGuinness, E. *Perception* **14**, 105-126 (1985).
16. Van Essen, D. C. *A. Res. Neurosci.* **2**, 227-263 (1979).
17. Movshon, J. A., Adelson, E. H., Gizzi, M. S. & Newsome, W. T. *Study Group on Pattern Recognition Mechanisms* (ed. Chagas, C., Gattass, R. & Gross, C. G.) (Vatican City, Pontificia Academia Scientiarum, 1984).
18. Albright, T. D. *J. Neurophysiol.* **52**, 1106-1130 (1984).
19. Rodman, H. R. & Albright, T. D. *Expl Brain Res.* **75**, 53-64 (1989).

ACKNOWLEDGEMENTS. We thank Nicolaus Bourbaki, Daniel Plummer, Terry Sejnowski, Diane Rogers-Ramachandran and Chandramani Ramachandran for discussion, and Bob Boynton for the use of his photometer. G.S. was supported by a Fight for Sight Summer Fellowship, T.D.A. by a grant from the National Eye Institute, and V.S.R. by a grant from the Airforce Office of Scientific Research.

Neuropeptide fusion of two motor-pattern generator circuits

Patsy S. Dickinson*, Carol Meccas* & Eve Marder†

*Biology Department, Bowdoin College, Brunswick, Maine 04011, USA

†Biology Department, Brandeis University, Waltham, Massachusetts 02254, USA

ANIMALS make many different movements as circumstances dictate. These movements often involve the coordination of several neural networks, the output of which can be changed by modulatory substances¹⁻⁴. Here we report that the neuropeptide red pigment concentrating hormone modulates the interactions between two rhythmic pattern-generating networks in the lobster stomatogastric nervous system. Red pigment concentrating hormone markedly enhances the amplitude of synaptic interactions between elements of two pattern-generating networks—the cardiac sac and the gastric mill. Consequently, two networks operating under some circumstances virtually independently can be fused into one functional unit operating differently from either of the two original networks. These results show how a neuropeptide can alter the functional configuration of a neural network so that widely disparate outputs can be produced by the same neurons.

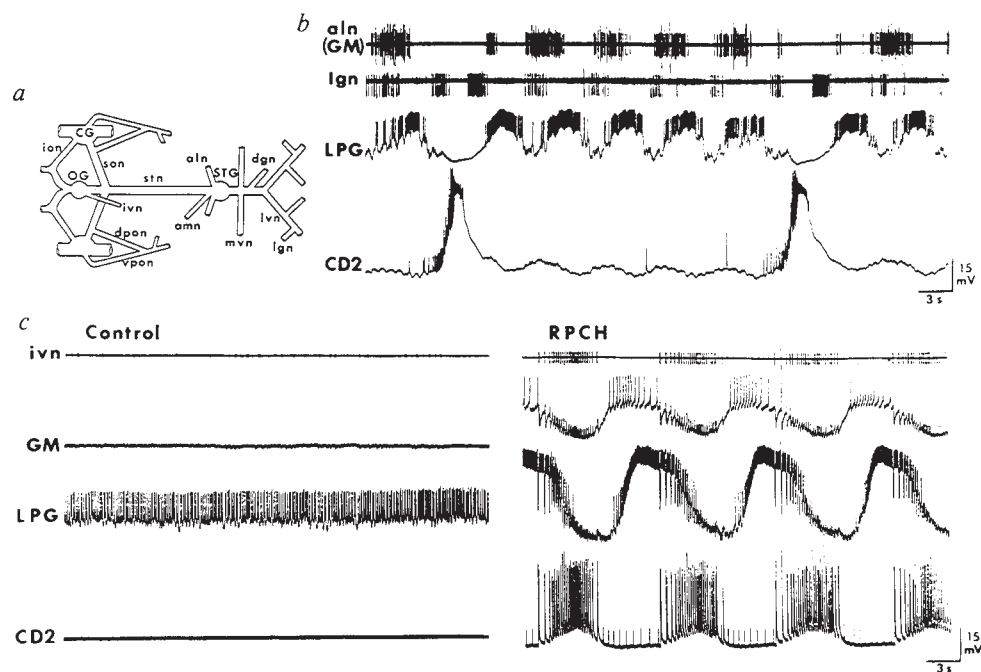
The stomatogastric nervous system of the lobster, *Panulirus interruptus*, consists of four ganglia (Fig. 1a) that together produce the four different rhythms of the foregut⁵. The two rhythms discussed here are the cardiac sac rhythm (shown as recordings of the cardiac dilator motor neuron 2 (CD2) and inferior ventricular nerve (ivn) in Fig. 1) and the gastric mill rhythm (shown as recordings of the lateral posterior gastric (LPG) and gastric mill (GM) neurons in Fig. 1). Figure 1b shows spontaneous activity in the cardiac sac and the gastric mill under conditions in which the four ganglia of the stomatogastric nervous system are left connected. The cardiac sac typically has a long period (20 s to several minutes; the record shown in Fig. 1b is one of the most rapid we have recorded), whereas the gastric mill typically has a shorter period of 5-10 s.

Red pigment concentrating hormone (RPCH) is found in processes throughout the neuropile of the stomatogastric ganglion (STG)^{6,7}. We have shown previously that RPCH strongly activates the cardiac sac rhythm in *P. interruptus*⁷. We now show that activation of the cardiac sac by RPCH is associated with massive changes in the patterns of activity in the gastric neurons. When inputs from the commissural ganglia are removed, the level of spontaneous activity in these two networks decreases, so that they are both generally silent, as shown in Fig. 1c, left. If the stomatogastric ganglion is superfused with RPCH under these conditions, rhythmic activity is once again recorded in the neurons of both the gastric mill and the cardiac sac networks (Fig. 1c, right). But the neurons of both networks are now active in a single rhythm. This is seen in Fig. 1c (right), in which GM, LPG, and CD2 are sequentially rhythmically active. This new rhythm, in which CD2 and gastric neurons are coordinately active, is typically slower (mean period, 28 s; $n = 37$) than the gastric rhythm (the RPCH rhythm in Fig. 1c (right) is among the fastest we have recorded).

Received 4 October 1989; accepted 11 January 1990.

1. Wallach, H. *On Perception* (Quadrangle, New York, 1976).
2. Adelson, E. H. & Movshon, J. A. *Nature* **300**, 523-525 (1982).
3. Hildreth, E. C. *The Measurement of Visual Motion* (MIT, Cambridge, 1984).
4. Ramachandran, V. S. & Anstis, S. M. *Scient. Am.* **254**, 102-109 (1986).
5. Ramachandran, V. S. & Cavanagh, P. *Vision Res.* **27**, 97-106 (1987).
6. Welch, L. *Nature* **337**, 734-736 (1989).
7. Murphy, B. J., Kowler, E. & Steinman, R. M. *Vision Res.* **15**, 1263-1268 (1975).
8. Beck, J. in *Human and Machine Vision II* (ed. Rosenfeld, A.) 1-12 (Academic, New York, 1986).
9. Metelli, F. *Scient. Am.* **230**, 91-95 (1974).
10. Nakayama, K., Shimojo, S. & Ramachandran, V. S. *Perception* (in the press).
11. Crick, F. H. C. *Nature* **337**, 129-132 (1989).
12. Ramachandran, V. S. in *Vision: Coding and Efficiency* (ed. Blakemore, C.) (Cambridge University Press, 1990).

FIG. 1 The stomatogastric nervous system of *P. interruptus*. *a*, Diagram of the stomatogastric nervous system. *b*, Gastric and cardiac sac rhythms in control preparation. Upper three traces, gastric rhythm; lower trace (CD2), cardiac sac rhythm. Recording methods as described in ref. 7. *c*, RPCH alters interactions of gastric and cardiac sac rhythms. Left traces, superior oesophageal nerves blocked with isotonic sucrose. Neither gastric nor cardiac sac pattern was rhythmically active. Right traces, STG superfused with 10^{-6} M RPCH. Cells in both patterns active in a single rhythm. (GM and LPG, normally gastric; ivn and CD2, cardiac.) Abbreviations: aln, anterior lateral nerve; amn, anterior median nerve; CD2, cardiac dilator neuron 2; CG, commissural ganglion; dgn, dorsal gastric nerve; dpon, dorsal posterior oesophageal nerve; GM, gastric mill neuron; ion, inferior oesophageal nerve; ivn, inferior ventricular nerve; lgn, lateral gastric nerve; LPG, lateral posterior gastric neuron; lvn, lateral ventricular nerve; mvn, median ventricular nerve; OG, oesophageal ganglion; son, superior oesophageal nerve; STG, stomatogastric ganglion; stn, stomatoga-



tric nerve; vpon, ventral posterior oesophageal nerve.

We were interested in determining the mechanism by which the neurons of the gastric rhythm fire coincidentally with those of the cardiac sac. One component of the cardiac sac rhythm, the IV neurons, fires spike bursts together with CD2⁸, but also makes strong monosynaptic connections with many of the neurons of the gastric mill and pyloric networks^{9,10}. We therefore studied the effects of bath application of RPCH on the amplitude of the synaptic potentials evoked by IV neuron activity in a variety of postsynaptic targets in the gastric rhythm. The IV neurons elicit large excitatory postsynaptic potentials (e.p.s.ps) in LPG neurons (Fig. 2, left). These e.p.s.ps are increased 1.5 to 4-fold ($n = 14$) in the presence of RPCH. Figure 2 (middle) shows inhibitory postsynaptic potentials (i.p.s.ps) evoked by the IV neurons in GM neurons. These i.p.s.ps are also enhanced 1.5 to 4-fold ($n = 23$) in the presence of RPCH. Figure 2 (right) shows that the e.p.s.ps evoked in CD2 by IV neurons stimulation are strongly (3 to 9-fold, $n = 19$) potentiated in RPCH.

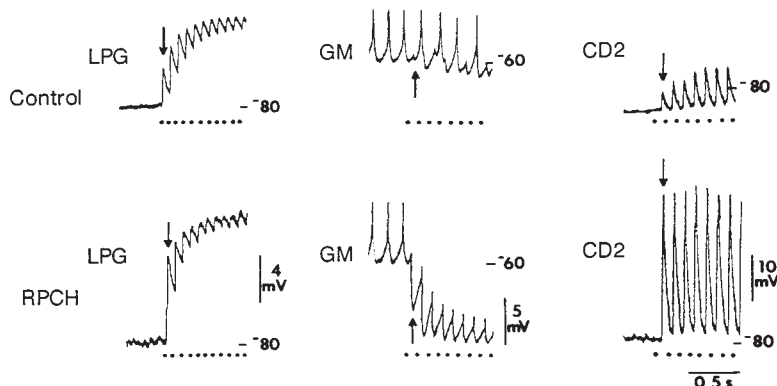
To determine if these increases in postsynaptic potential (p.s.p.) amplitudes are due to a change in postsynaptic membrane potential or input impedance, we placed two electrodes in the postsynaptic neurons, and measured p.s.p. amplitude as a function of membrane potential. Additionally, voltage-current curves were measured for each postsynaptic neuron. The RPCH-induced increases in p.s.p. amplitude were seen at all levels of

membrane potential tested, and the peptide produced no apparent changes in the measured or extrapolated reversal potentials of the p.s.ps (Fig. 3*a, c, d*). Figure 3*b* shows a voltage-current curve for the LPG neuron in control saline and in RPCH. These plots show that the increase in e.p.s.p. amplitude cannot be accounted for by a change in the input impedance of the cell. Similar results were found in four to six experiments on GM, LPG and CD2 neurons.

The data presented here indicate that the motor pattern found in the presence of RPCH, in which elements of the gastric rhythm and cardiac sac rhythm are jointly active, can be simply explained on the basis of a massive increase in the strength of the synaptic interactions between a set of neurons that drive the cardiac sac rhythm and the follower neurons that normally form part of the gastric mill (and also pyloric) pattern generators, as illustrated in Fig. 4. The data presented here do not allow us to ascertain unambiguously whether the increase in the amplitude of the synaptic potentials from the IV neurons seen in the presence of RPCH is due to an increase in transmitter release from the synaptic terminals of the IV neurons or due to some postsynaptic mechanism, not ruled out by the lack of change in the voltage-current relations.

Hoooper and Moulins¹¹ recently showed that a neuron originally classified as part of the pyloric network can switch between the cardiac sac rhythm and the pyloric rhythm. Meyrand

FIG. 2 The p.s.ps from the IV neurons of the cardiac sac pattern generator increase in amplitude when the STG is superfused with 10^{-6} M RPCH. Superior oesophageal nerves blocked with sucrose to decrease spontaneous activity. The ivn (containing axons of IV neurons) was stimulated at 10 Hz and the resulting p.s.ps recorded. Excitatory p.s.ps in LPG and CD2, and inhibitory p.s.ps in GM were potentiated. Dots, ivn stimulation; arrows, peak of first p.s.p. in each trace.



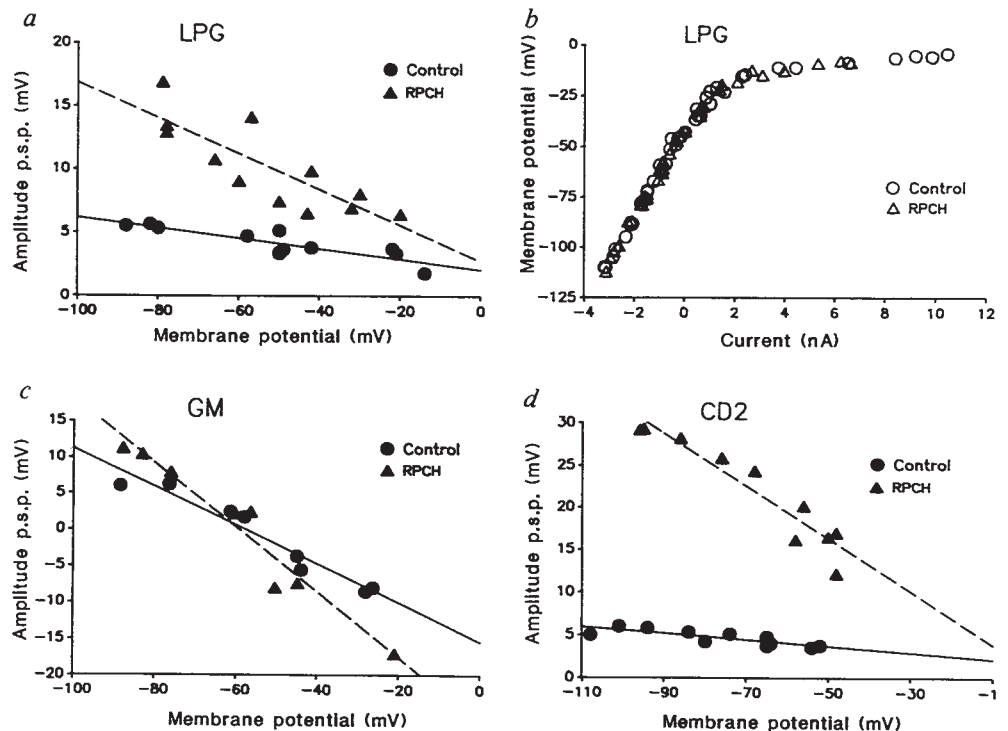


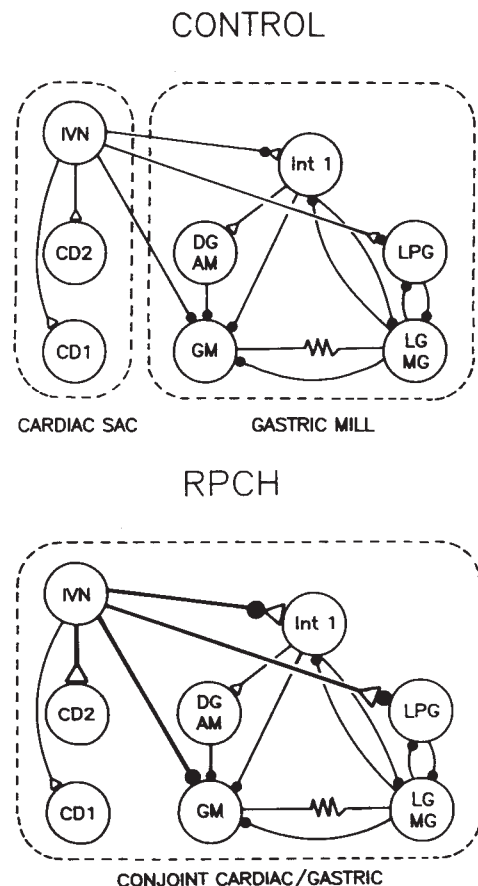
FIG. 3 The increase in amplitude of IV neuron p.s.p.s cannot be accounted for by changes in membrane potential or reversal potential. The p.s.p. amplitude plotted as function of postsynaptic membrane potential for LPG (a), GM (c) and CD2 (d). Amplitude was greater in the presence of RPCH (triangles) at all membrane potentials; apparent reversal potentials were unchanged. Input resistance, measured in current-clamp in cell body, was also unchanged, as seen for LPG in b. Similar curves for GM and CD2 are not shown.

FIG. 4 Diagrammatic representation of the changes in p.s.p. amplitude that seemed to be responsible for the fusion of the gastric mill and cardiac sac motor patterns. Control, established circuits⁵ for gastric and cardiac-sac patterns as well as the synapses of the IV neurons, through which the two patterns normally interact. RPCH, the IV neuron synapses dominate both patterns and lead to establishment of single conjoint gastric-cardiac-sac pattern. ●, Inhibitory synapses; △, excitatory synapses; ~, electrical coupling. Abbreviations as in Fig. 1; AM, anterior median neuron; CD1, cardiac dilator neuron 1; DG, dorsal gastric neuron; Int 1, interneuron 1; LG, lateral gastric neuron; MG, median gastric neuron; IVN, inferior ventricular neurons.

*et al.*¹² and Weimann *et al.*¹³ have recently reported that many crab STG neurons can switch between pyloric and gastric patterns. These reports, together with the one presented here, indicate that neurons identified primarily on the basis of their participation in one motor pattern can participate in the generation of several patterns. This calls into question some of the assumptions underlying identification of neurons on functional grounds.

There are significant differences between the neuronal 'switches' previously described¹¹⁻¹³ and this study. In the former, individual neurons or neuronal groups "change their allegiance" between different ongoing rhythmic motor patterns. In the study described here, RPCH converts two independent motor patterns of significantly different periods into a single conjoint rhythm that is appreciably different from either of the initial rhythms. We find it interesting that in the study by Hooper and Moulins¹¹, activation of a sensory input elicits the switch of the ventricular dilator neuron by suppressing its plateau properties. By contrast, the merging of the cardiac sac and gastric rhythms described here is induced by RPCH modification of the efficacy of defined synaptic contacts.

With these data we can begin to understand how a neural network composed of synaptically connected neurons can be not only modulated, as previously found¹⁻⁴, but massively reshaped, so that entirely different behaviours can be built from the same individual components. Specifically, the mechanism underlying the reconfiguration described in this report is peptide-induced potentiation of one series of synaptic potentials that strongly coordinate activity in postsynaptic neurons ordinarily functioning in separate neural networks. Because



RPCH is present in defined inputs to the STG⁷, exogenous application of RPCH is likely to mimic activation of the RPCH-containing input fibres. RPCH is only one of many modulatory substances found in inputs to the STG¹⁴. We have therefore only begun to specify the full constellation of behaviours possible for a single network of related neural circuits. □

Received 15 December 1989; accepted 3 January 1990.

1. Marder, E. & Hooper, S. L. in *Model Neural Circuits and Behavior* (ed. Selverston, A. I.) 319–337 (Plenum, New York, 1985).
2. Harris-Warrick, R. M. in *Neural Control of Rhythmic Movements in Vertebrates* (eds Cohen, A. V., Rossignol, S. & Grillner, S.) 285–331 (Wiley, New York, 1988).
3. Dickinson, P. S. *Semin. Neurosci.* **1**, 15–24 (1989).
4. Marder, E. *Semin. Neurosci.* **1**, 3–4 (1989).
5. Selverston, A. I. & Moulins, M. (eds) *The Crustacean Stomatogastric System* (Springer, Berlin, 1987).
6. Nusbaum, M. P. & Marder, E. *J. exp. Biol.* **135**, 165–181 (1988).
7. Dickinson, P. S. & Marder, E. *J. Neurophysiol.* **61**, 833–844 (1989).
8. Moulins, M. & Vedel, J. P. *J. Physiol., Paris* **73**, 471–510 (1977).
9. Sigvardt, K. A. & Mulloney, B. *J. exp. Biol.* **97**, 153–168 (1982).
10. Claiborne, B. J. & Selverston, A. I. *J. comp. Physiol.* **A154**, 27–32 (1984).
11. Hooper, S. L. & Moulins, M. *Science* **244**, 1587–1589 (1989).
12. Meyrand, P., Weimann, J. M. & Marder, E. *Soc. Neurosci. Abstr.* **14**, (1988).
13. Weimann, J. M., Meyrand, P. & Marder, E. in *Crustacean Systems in Neurobiology* (ed. Wiese, K.) (Birkhäuser, Basel, in the press).
14. Marder, E. in *The Crustacean Stomatogastric System* (eds Selverston, A. I. & Moulins, M.) 263–300 (Springer, Berlin, 1987).

ACKNOWLEDGEMENTS. We are indebted to Dr Scott Hooper, Dr Pierre Meyrand and Mr James Weimann for discussions. We thank P. Cason and M. B. O'Neil for help with the figures. The research was supported by the NSF and the NIH.

Segmental and developmental regulation of a presumptive T-cell oncogene in the central nervous system

J. M. Greenberg*, T. Boehm, M. V. Sofroniew†, R. J. Keynes†, S. C. Barton‡, M. L. Norris‡, M. A. Surani‡, M-G. Spillantini & T. H. Rabbitts

Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

† Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB3 2DY, UK

‡ Department of Molecular Embryology, AFRC Institute of Animal Physiology and Genetics Research, Babraham, Cambridge CB2 4AT, UK

ALTHOUGH most proto-oncogenes such as *c-myc* are involved in cell proliferation, being expressed in a wide range of tissues as well as in progenitors of transformed cells¹, others may normally function in cellular differentiation². We now report on a gene on human chromosome 11, at the junction of the T-cell tumour-associated chromosomal translocation t(11;14) (p15;q11) and known as the *11p15* gene (ref. 3) or *Ttg* (ref. 4), which is believed to be involved in the pathogenesis of the tumour. It has two transcriptional promoters (both retained by the translocated allele⁵) and is expressed in tumour cells with neuro-endocrine properties, suggesting that normal expression may occur in nerve cells. Using fusion constructs of one *11p15* promoter and *lacZ* in transgenic mice, we found that the gene is expressed in a segment-specific manner in rhombomeres^{6–13} of the developing mouse hindbrain. During subsequent development, the gene is more widely expressed, again in precisely defined regional patterns, but in post-mitotic neurons confined to the central nervous system. Thus, this presumptive T-cell oncogene is both developmentally regulated and segmentally restricted in a tissue different from that in which the original tumour arose.

To establish the expression of the *11p15* gene *in vivo* we used a constructed plasmid, introduced into the germ line of mice, in which the bacterial *lacZ* gene is under the transcriptional control of one *11p15* gene promoter (Fig. 1). This construct determines the expression of bacterial β -galactosidase in all tissues where the *11p15* gene promoter 1 is active as the presence of this exogenous enzyme renders cells blue when appropriately stained¹⁴.

Embryos of various ages from five founders were analysed for β -galactosidase activity (Fig. 1 legend). Similar patterns of expression were observed in progeny of all founders at different stages of development. Stages from pre-implantation embryos were studied. Expression starts on day 9 of gestation (E9) and, thereafter, is restricted to the developing nervous system. From days E9 to E11, β -galactosidase expression is most prominent in the hindbrain, in a segmental pattern involving rhombomeres 2 and 4 (Figs 2a, b and 3a). On the basis of the size and position of the cell bodies, and the presence of β -galactosidase in axons extending into the adjacent branchial arches (Figs 2a, b and 3a), we concluded that the stained structures within the rhombomeres are postmitotic, branchiomotor neurons, respectively trigeminal (V) and facial (VII)¹⁰. In addition, the pattern of staining evolves to include similar cells within rhombomeres r3 and r5, matching the temporal and spatial differentiation of these branchiomotor neurons¹⁰. It is also possible that some glia may express the transgene. In view of the initial expression of the *11p15* gene within rhombomeres (Fig. 2a, b) and its identification at a chromosome translocation in a T-cell tumour, we have called this gene the rhombotin gene.

Although the most prominent initial transgene expression is observed in rhombomeres, some staining of other central nervous system (CNS) structures can be seen by E10 (Fig. 2a); this includes scattered large neurons in the spinal cord ventral horn as well as cell populations in the ventral forebrain and retina. Expression is restricted to the CNS, and neural crest derivatives, such as the sensory ganglia, are not stained (Fig. 2a). During days E15 and E16, gene expression becomes apparent in the developing cerebral neocortex, labelling postmitotic cells along its pial border. The ventricular zone is negative. Staining is also prominent in various cell groups in the ventral forebrain and hindbrain.

Examination of neonatal transgenic brains (Fig. 2c) shows that β -galactosidase is expressed in most regions of the cerebral cortex at P0–P2, labelling postmitotic cells primarily in the middle layers (Fig. 4a). The olfactory bulbs and hippocampus also show staining in specific cell layers and staining is prominent in scattered cell groups throughout the ventral forebrain, midbrain (Fig. 2c) and medulla oblongata (Figs. 2c and 3b). In the cerebellar cortex, a striking pattern of longitudinal stripes

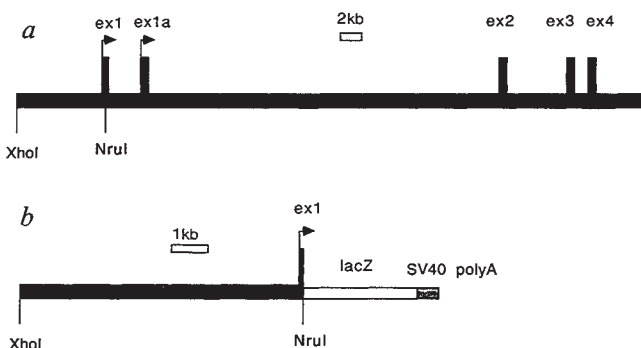
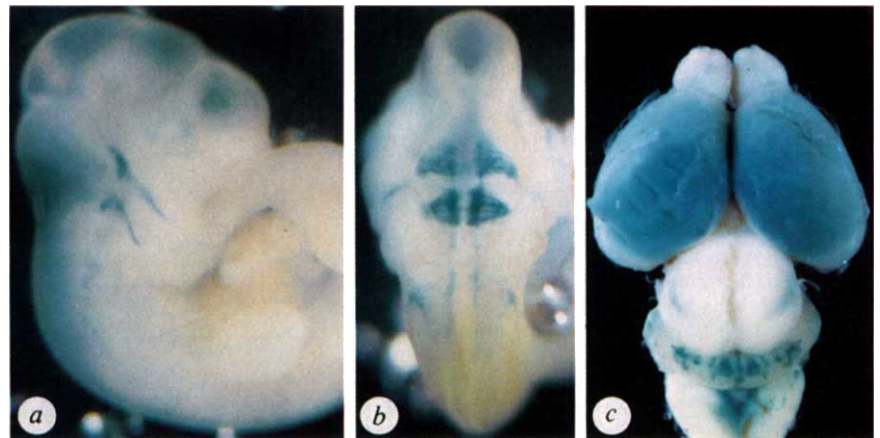


FIG. 1 Structure of the *11p15-lacZ* fusion transgene. *a*, Organization of the *11p15* gene from human chromosome 11 showing the five exons of the gene and the alternative promoters 1 and 1a⁵. The relevant restriction enzyme sites used for the subcloning are shown. *b*, The structure of the *11p15-lacZ* fusion construct. The *XhoI*–*NruI* fragment containing the promoter 1 and about 8 kb of genomic DNA from the human *11p15* gene⁵ was joined to the bacterial *lacZ* gene segment isolated from the pCH101 clone²⁴ in the vector pBlueScript. The *NruI* site is within exon 1 of the *11p15* gene⁵ and occurs before the first donor splice site of this gene. The gel-purified fragment of the construct was isolated, for micro-injection into mouse eggs, after digestion with *XhoI*. Procedures for the generation of transgenic mice and the assay for β -galactosidase have been described¹⁴. Five independent transgenic founders, amongst 24 progeny, were studied after transfer of over 200 fertilized (C57B × CBA)F₁ eggs, micro-injected with the gene construct. The five founder mice carried low copy numbers of the transgene.

* Present address: University of Minnesota Medical School, Harvard Street at E. River Road, Minneapolis, USA.

FIG. 2 Expression of the 11p15-lacZ fusion gene construct in the mouse brain. *a*, Lateral view of a whole transgenic embryo at day E11.5 showing prominent blue staining resulting from β -galactosidase activity¹⁴ in trigeminal (V) and facial (VII) branchiomotor neurons and their axons, extending into the first and second branchial arches respectively. Some faintly stained cells are present in the lower brainstem and spinal cord. *b*, Dorsal view of the day E11.5 embryo (shown in 2*a*) showing β -galactosidase activity in developing branchiomotor neurons V and VII in rhombomeres 2 and 4. A small number of faintly stained cells are present along the midline in rhombomeres 3, 5, 7 and 8. *c*, Dorsal view of a whole brain of a neonatal (P1) transgenic mouse. Prominent blue staining showing the presence of β -galactosidase is visible in cerebral neocortex (forebrain), superior colliculus (midbrain), cerebellum (anterior hindbrain) where expression is evident as prominent longitudinal stripes and



dorsomedial medulla oblongata (caudal hindbrain).

is created by bands of cells which alternately express the transgene (Figs. 2*c* and 3*b*). The banding is reminiscent of the longitudinal organization of fibre connections in the cerebellum¹⁵, but whether the two are related is not yet known. No transgene expression was noted in structures outside the CNS and, on the basis of their size and morphology, most labelled cells, but perhaps not all, appear to be neuronal.

The pattern of CNS expression of the fusion gene in the progeny of five independent founder mice indicates that transgene expression is not artefactual. This conclusion was verified by assessing the expression of the endogenous rhombotin gene in normal mice by *in situ* hybridization (Fig. 4) using a mouse complementary DNA clone⁵. Sections through newborn non-transgenic mouse brains were hybridized to the cDNA probes to detect endogenous rhombotin messenger RNA. Both sense (Fig. 4*b*) and antisense (Fig. 4*d*) probes were used but only the antisense probe hybridized. The specificity of transgene expression was compared in detail with this hybridization signal (Fig. 4). The pattern of mRNA *in situ* hybridization largely paralleled the expression of the transgenic fusion gene in most CNS regions, including the cerebellum, hippocampal formation, cerebral cortex and olfactory bulbs. Figure 4 shows similar patterns of transgenic expression and endogenous rhombotin gene expression in the middle layers of the cortex and in the sub-fields CA1 and CA3 of the hippocampus. Interestingly, granule cells of the dentate gyrus and cells in the stratum radiatum and the stratum oriens of the hippocampus appear negative by both procedures, confirming the specificity and equivalence of the transgene and endogenous gene expression.

The transgenic construct used in these experiments has about 8 kilobases (kb) of genomic DNA, including the transcription promoter. The specific developmental and segmental CNS expression pattern described here is guided entirely by sequences within this 8 kb, which are, therefore, sufficient to facilitate this regulated expression. There are, however, some CNS regions (for example, the caudate putamen) where endogenous rhombotin gene expression is observed by *in situ* hybridization but not transgenic expression (data not shown). These differences presumably are because the transgene lacks other necessary elements.

These data show that the rhombotin gene, which was originally identified by a chromosomal abnormality in a T-cell tumour³, is both developmentally and segmentally expressed in mouse brain. This gene is a presumptive proto-oncogene, by analogy with *c-myc*, *bcl-2* and *c-abl*, which are found at chromosomal translocations in Burkitt's lymphoma, follicular lymphoma and chronic myelogenous leukaemia, respectively, (see ref. 16 for review). Although some other oncogenes are expressed in the CNS¹⁷⁻²⁰, their expression is not restricted to neurons. The function of the rhombotin gene in the developing CNS is

not yet known, but the comparison between its segmental pattern of expression and the expression of certain mouse homoeobox genes^{9,12,13,21,22} and the zinc-finger gene *Krox-20*¹¹ is interesting. At E12, the anterior boundaries of expression of the Hox-2 gene cluster align with alternate rhombomere boundaries¹³ and *Krox-20* is transcribed in alternate rhombomeres r3 and r5. However, the rhombotin gene is not a homoeobox gene^{4,5}, and our data indicate that the rhombotin protein is not involved in segment determination. Instead, its expression in postmitotic branchiomotor neurons suggests that it may play a role in determining this specific phenotype. Its initial appearance in postmitotic

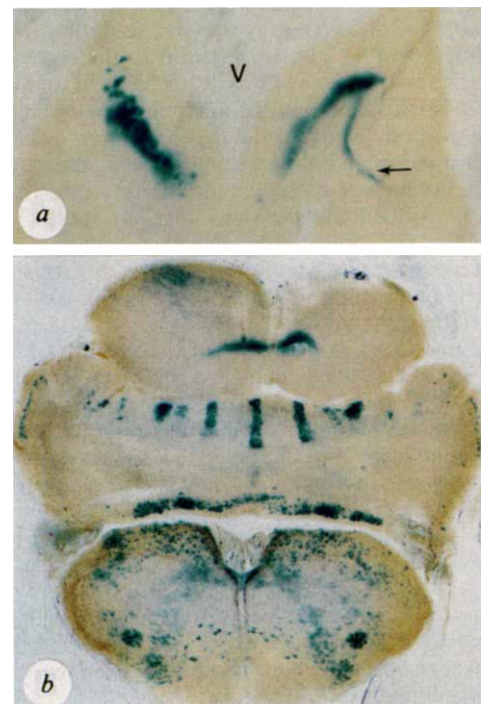
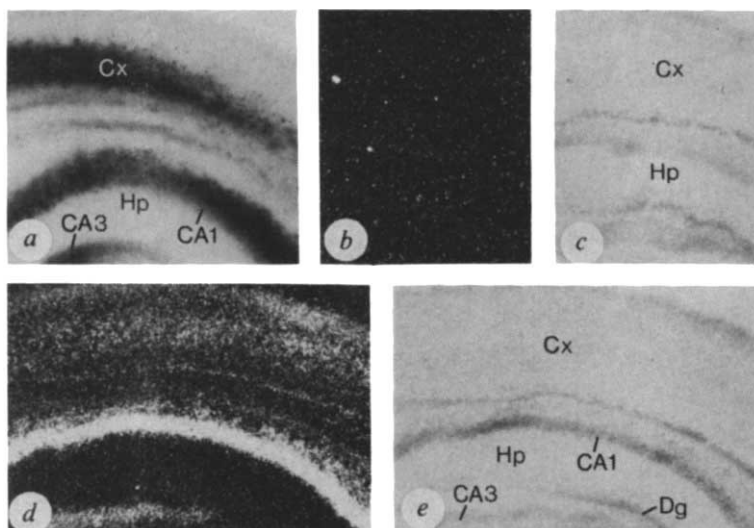


FIG. 3 Expression pattern of the 11p15-lacZ fusion gene construct in sections of transgenic brains. *a*, Transverse section through the hindbrain of a transgenic embryonic mouse at day E11.5. Blue staining indicating β -galactosidase activity is present in large cells comprising branchiomotor neurons of rhombomere 2. Axons from these neurons leave the CNS, and pass medially to the trigeminal ganglion into the first branchial arch (arrow). Note that the trigeminal ganglion and periventricular zone are unstained. V, 4th ventricle. *b*, Transverse section through the brain stem and cerebellum of a transgenic neonatal mouse at day P1. Blue-stained cells containing β -galactosidase activity are present in the superior colliculus and scattered in several regions of the medulla oblongata. In the cerebellum, stained cells are organized into a pattern of longitudinally orientated stripes.

FIG. 4 Comparison of transgenic β -galactosidase expression and cDNA *in situ* hybridization of non-transgenic mouse brain. Transverse sections through the cerebral cortex and hippocampal formation of a transgenic mouse or a normal mouse, at day P1, to compare β -galactosidase expression with *in situ* hybridization signal (sense probe in *b*, anti-sense probe in *d*). The *in situ* probes were derived from a mouse cDNA clone (N2AIII) and correspond to nucleotides 97–567 (the latter position is at a *Ball* restriction site) of the previously described clone pN2A⁵. *In situ* hybridization was carried out as previously described^{25,26} using prime-cut ³⁵S-labelled probes derived from cDNA cloned in both orientations in M13. *a*, Section of transgenic brain showing β -galactosidase activity (dark staining) in postmitotic cells in the middle layers of the cerebral cortex (Cx) and in CA1 and CA3 subdivisions of the hippocampus (Hp). *b*, Dark-field photomicrograph of a transverse section through a similar level of the forebrain of a non-transgenic mouse, showing *in situ* hybridization using the cDNA sense probe. No specific hybridization is evident. *c*, Bright-field photomicrograph of precisely the same area depicted in *b*, indicating the location of the hippocampus and cortex. *d*, Dark-field photomicrograph of a section close to that shown in *b*, showing *in situ* hybridization with the cDNA anti-sense probe. Binding is prominent in CA1 and CA3 subdivisions of the hippocampus and in middle layers of the overlying cortex. *e*, Bright-field photomicrograph of the same area shown in *d*, indicating the location of



the hippocampus, CA1, CA3, cortex and the dentate gyrus (Dg). Note that the dentate gyrus shows no hybridization of the cDNA anti-sense probe in *d*, and no expression of β -galactosidase.

neurons indicates that it is involved in differentiation rather than cell proliferation.

Most known proto-oncogenes seem to be involved in cell proliferation but it has been suggested that there is a second class, which probably overlaps with the first, which normally functions in differentiation². In these cases, ectopic expression may be the key feature in tumorigenesis. For example, the *int-2* proto-oncogene (discovered by pro-viral insertion) is normally found in neuroepithelial cells adjacent to the otocyst, as well as in other sites outside the CNS²⁰. A variety of tumour-associated chromosomal abnormalities are found in lymphoid tumours and some of these involve cellular counterparts of known viral oncogenes involved in proliferation. However, most of the cases studied do not have retroviral counterparts¹⁶. One possibility is that these newly identified loci carry stage-specific differentiation genes, whose capacity for involvement in tumorigenesis is restricted. Because they are more restricted targets, they would be associated with somewhat rare translocations such as those found in T-lineage tumours²³.

The rhombotin gene represents the first example of the second class of proto-oncogene, activated by the generation of a chromosomal abnormality. Under normal conditions, this gene is developmentally and segmentally expressed in the CNS. Various studies have, thus far, failed to show that normal T-cells express the gene from either promoter 1 or 1a (ref. 5); these included the examination of RNA made from mouse thymus of various maturities (although we cannot rule out the possibility that a T-cell subset expresses the gene). It is, therefore, probable that the gene is rendered transcriptionally active by chromosomal translocation, in the pre-T-cell tumour, which, in this case, leads to leukaemia development. Other T-cell tumour-associated chromosomal abnormalities may also reveal developmentally regulated proto-oncogenes. □

Received 24 November 1989; accepted 19 January 1990.

- Kelly, K., Cochran, B., Stiles, C. & Leder, P. *Cell* **35**, 603–610 (1983).
- Knudson, A. G. *Rev. Genet.* **20**, 231–251 (1986).
- Boehm, T. *et al. EMBO J.* **7**, 385–394 (1988).
- McGuire, E. A. *et al. Molec. cell. Biol.* **9**, 2124–2132 (1989).
- Boehm, T. *et al. EMBO J.* (in the press).
- von Baer, K. E. *Über die Entwicklungsgeschichte der Thiere* (Konigsberg, 1828).
- Orr, H. J. *Morph.* **1**, 311–372, (1887).
- McClure, C. F. W. *J. Morph.* **4**, 35–56 (1890).
- Gaunt, S. J. *Development* **101**, 51–60 (1987).
- Lumsden, A. & Keynes, R. *Nature* **337**, 424–428 (1989).
- Wilkinson, D. G. *et al. Nature* **337**, 461–464 (1989).
- Murphy, P., Davidson, D. R. & Hill, R. E. *Nature* **341**, 156–159 (1989).
- Wilkinson, D. G., Bhatt, S., Cook, M., Boncinelli, E. & Krumlauf, R. *Nature* **337**, 405–409 (1989).
- Allen, N. D. *et al. Nature* **333**, 852–855 (1988).
- Groenewegen, H. J. & Voogd, J. *J. Comp. Neurol.* **174**, 417–488 (1977).
- Boehm, T. & Rabbitts, T. H. *Eur. J. Biochem.* **185**, 1–17 (1989).
- Peters, G., Brookes, S., Smith, R. & Dickson, C. *Cell* **33**, 364–377 (1983).
- Gertler, F. B., Bennett, R. L., Clark, M. J. & Hoffman, F. M. *Cell* **58**, 103–113 (1989).
- Klein, R., Parada, L. F., Coulter, F. & Barbacid, M. *EMBO J.* **8**, 3701–3709 (1989).
- Wilkinson, D. G., Peters, G., Dickson, C. & McMahon, A. P. *EMBO J.* **7**, 691–695 (1988).
- Kappen, C., Schughart, K. & Ruddle, F. H. *Ann. N.Y. Acad. Sci.* **VI**, 243–252 (1989).
- Fainsod, A., Awgulewitsch, A. & Ruddle, F. H. *Dev. Biol.* **124**, 125–133 (1987).
- Rabbitts, T. H., Boehm, T. & Mengle-Gaw, L. *Trends Genet.* **4**, 300–304 (1988).
- Hall, C. V., Jacob, P. E., Ringold, G. M. & Lee, F. J. *J. molec. appl. Genet.* **2**, 101–109 (1983).
- Goedert, M. *EMBO J.* **6**, 3627–3632 (1987).
- Spillanti, M. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 8555–8559 (1989).

ACKNOWLEDGEMENTS. We thank I. Lavenir, A. Forster and S. Stevens for technical assistance, J. Bashford for photography, and M. Goedert for help and advice on techniques and the manuscript. J.M.G. was the recipient of a Fellowship of the Leukemia Society of America.

Impairment of endothelium-dependent arterial relaxation by lysolecithin in modified low-density lipoproteins

Kiyotaka Kugiyama*, Scott A. Kerns*, Joel D. Morrisett†, Robert Roberts* & Philip D. Henry*‡§

Departments of * Medicine, † Biochemistry, and ‡ Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, Texas 77030, USA

ATHEROSCLEROSIS in animals and humans is associated with an unresponsiveness of arteries and arterioles to endothelium-dependent vasodilators^{1–3}—agents acting on smooth muscle indirectly by stimulating the release from endothelial cells of a vasodilator principle (endothelium-derived relaxing factor)⁴. Altered vasomotor regulation in atherosclerosis could partly reflect an injurious action of abnormal lipoproteins on endothelium^{5,6}. Recently, 'cell-modified' or 'oxidized' low-density lipoprotein (EC-LDL) has received increasing attention because of its potential cytotoxic and atherogenic properties⁷. We report here that arteries exposed to EC-LDL *in vitro* show an endothelium-dependent

§ To whom correspondence should be addressed.

vasoregulatory impairment closely resembling that of atherosclerotic arteries. Our results indicate that transfer of lysolecithin from EC-LDL to endothelial membranes produces a selective unresponsiveness to receptor-regulated endothelium-dependent vasodilators.

Atherosclerotic vessels show a characteristic response to vasodilators¹⁻³. They have a decreased responsiveness to endothelium-dependent vasodilators (EDVs) that act through endothelial surface receptors (acetylcholine, ATP, substance P), whereas their responsiveness to EDVs that bypass receptor-dependent membrane regulation, the calcium ionophores (A23187, ionomycin), is partly or completely preserved. In addition, there is complete preservation of responsiveness to agents acting directly on smooth muscle, such as nitroglycerin. To assess the contribution of lipoproteins to this vasomotor syndrome, we characterized the vasodilator response of isolated nonatherosclerotic rabbit aortas before and after their exposure to selected lipoproteins and lipids. We obtained preparations of native LDL (N-LDL) and of LDL modified by incubation with endothelial cells in culture (EC-LDL) by following published protocols⁸⁻¹⁰. After incubation with endothelial cells, LDL showed characteristic changes including increased electrophoretic mobility, accumulation of thiobarbituric acid-reactive substances (TBARS), and substantial degradation of lecithin to lysolecithin⁸⁻¹⁰. We re-isolated lipoprotein preparations before use to free them of cell-culture contaminants and preparative additives, such as antioxidants or albumin. The concentrations of lipoproteins in which we incubated arteries corresponded to estimated concentrations of oxidized LDL in arterial walls¹¹. After incubation with EC-LDL, but not N-LDL, a precontracted artery became refractory to the relaxing effects of acetylcholine, although nitroglycerin evoked complete relaxation (Fig. 1). Results of experiments with isolated arteries are summarized in Table 1 and Fig. 2. Incubations with EC-LDL substantially suppressed relaxations in response to acetylcholine, whereas N-LDL had no effect (Table 1). When we exposed N-LDL to cultured endothelial cells in the presence of the antioxidant butylated hydroxytoluene (BHT), changes in 'oxidative modification' were minimized and at the same time effects on endothelial function were markedly attenuated (Table 1). We tested fixed concentrations of vasodilators after incubations with varying lipoprotein concentrations (Fig. 2). Acetylcholine (1 μ M) was already inhibited with the lowest EC-LDL concentration, whereas responsiveness to A23187 (0.1 μ M) was

completely preserved up to 50 μ g protein ml⁻¹ (Fig. 2). With the highest EC-LDL concentration tested (100 μ g protein ml⁻¹), relaxations in response to acetylcholine (1 μ M), substance P (1 nM; not plotted), ATP (100 μ M; not plotted) and A23187 (0.1 μ M) were $7 \pm 2\%$, $26 \pm 3\%$, $39 \pm 4\%$, and $51 \pm 12\%$ (six to nine experiments for each agonist; Fig. 2). We performed dose-response experiments with vasodilators, and determined the vasodilator concentrations required to reduce phenylephrine (1 μ M)-induced constrictor tone by 20% (EC₂₀, an index of sensitivity) and 50% (EC₅₀) before and after incubation with EC-LDL (100 μ g protein ml⁻¹). Responses to acetylcholine after incubation were too small to permit determination of EC₂₀ and EC₅₀ values. With A23187, EC₂₀ was unchanged (3.2 ± 0.2 nM; $n = 7$), but EC₅₀ increased more than sixfold (from 5.0 ± 0.4 to 32.3 ± 4.0 nM; $n = 7$). With nitroglycerin, EC₂₀ (13 ± 0.5 nM; $n = 6$) and EC₅₀ (32 ± 2 nM; $n = 6$) values were nearly identical before and after incubation. Responses were not influenced by the presence during incubations of deferoxamine (10 μ M), indomethacin (10 μ M), superoxide dismutase (150 U ml⁻¹, Sigma S-4761), catalase (800 U ml⁻¹, Sigma C-40), and L-arginine (100 μ M) (five to seven experiments with each agent).

TABLE 1 Effects of LDL preparations on arterial relaxation and echinocyte formation

	Arterial relaxation* (% of pre-incubation control)	Echinocytes† (% of total RBC)
N-LDL	$98 \pm 2\%$	$<1\%$
C-LDL	$75 \pm 4\%$	$7 \pm 1\%$
EC-LDL	$7 \pm 2\%$	$40 \pm 4\%$
BHT-EC-LDL	$59 \pm 5\%$	$11 \pm 2\%$
A-EC-LDL	$72 \pm 5\%$	$<1\%$
N-LDL lipid	$96 \pm 2\%$	$<1\%$
EC-LDL lipid	$1 \pm 1\%$	$37 \pm 6\%$
LPC	$24 \pm 6\%$	$25 \pm 2\%$

* Aortic strips contracted with 1 μ M phenylephrine were relaxed with 1 μ M acetylcholine before and after 120-min exposure to preparations of LDL or lipid. Before incubation, percentage reduction in tone in response to acetylcholine was similar in the different groups with mean values ranging between 70 ± 2 and $74 \pm 4\%$. Preparations of LDL were obtained as described previously⁸⁻¹⁰ and re-isolated by ultracentrifugation before use. After re-isolation, all LDL preparations contained low levels of TBARS (<0.1 nmol malondialdehyde equivalents (MDA) per mg LDL protein). Abbreviations: N-LDL, native LDL; C-LDL and EC-LDL, N-LDL incubated for 24 h in culture dishes without (c, control) or with (EC) rabbit aortic endothelial cells grown to confluence. EC-LDL used before re-isolation ($n = 4$) and containing 5.3 ± 0.3 nmol MDA ml⁻¹ culture medium exerted effects that were very similar to those of re-isolated EC-LDL. Preparations of copper-oxidized LDL⁸ ($n = 7$) likewise mimicked effects of EC-LDL (results not shown). BHT-EC-LDL, same as EC-LDL, but medium contained 20 μ M butylated hydroxytoluene (antioxidant). A-EC-LDL, EC-LDL after incubation for 24 h with 100-fold excess of defatted albumin (0.15 mg LDL protein per 15 mg albumin); storage for 24 h under identical conditions but without albumin had no effect on EC-LDL activity; N-LDL lipid and EC-LDL lipid, lipid was extracted from the lipoproteins with chloroform-methanol (2:1 (V/V)) under precautions aimed at preventing oxidation (nitrogenation and fresh distillation of solvents containing 20 μ M BHT; temperature, 0–2 °C) and dispersed by sonication. LPC, palmitoyl-lysophosphatidylcholine. Incubations of arteries refer to 2-h exposures to preparations of lipoprotein (100 μ g LDL protein per ml Krebs buffer), dispersed lipid extracts (100 μ g cholesterol per ml Krebs buffer), or synthetic palmitoyl lysolecithin (10 nmol per ml Krebs buffer; for comparison, incubations with EC-LDL and EC-LDL lipid produced mean total lysolecithin concentrations of 59 ± 4 and 62 ± 5 nmol per ml buffer). Values for each type of LDL or lipid preparation represent means $\pm$ s.e.m. of six to nine.

† Fresh human erythrocytes (RBC) washed three times with iso-osmotic saline were incubated for 120 min at 37 °C in buffer (150 mM NaCl, 20 mM Tris-Cl, 1 mM CaCl₂, 10 mM glucose) containing different LDL or lipid preparations as described for artery incubations. The percentage of RBC converted to echinocytes (stages II–III, Bessis nomenclature¹⁷) was estimated microscopically. Some cell samples were also examined by electron microscopy.

‡ $P < 0.01$ compared with EC-LDL or EC-LDL lipid, respectively (unpaired t or Mann-Whitney tests). Values for both artery and red-cell experiments are means $\pm$ s.e.m. of six to nine independent experiments.

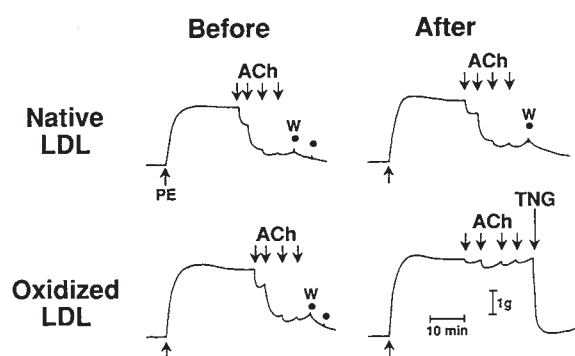


FIG. 1 Rabbit aortic strips mounted for the recording of isometric tension were equilibrated in oxygenated Krebs buffer at 37 °C and contracted with 1 μ M phenylephrine (PE)^{1,2}. Relaxation in response to graded concentrations of acetylcholine (ACh; 0.1, 0.5, 1.0 and 5 μ M) was determined before and after incubation with N-LDL or EC-LDL (both 100 μ g protein per ml buffer). All incubations were performed in the presence of normoxia ($pO_2 \approx 100$ mm Hg) and the antioxidant BHT (20 μ M) to prevent intra-experimental LDL degradation. After 120 min, the arteries were washed with Krebs buffer (W) to remove LDL and BHT, and dose-response experiments were repeated. EC-LDL, but not N-LDL, suppresses responses to ACh, but 1 μ M nitroglycerin (TNG) remains potent.

At the end of the experiments, we fixed arterial strips in 1% glutaraldehyde and prepared them for electron microscopic examination¹. After exposure to either N-LDL or EC-LDL, arteries showed a preserved endothelial cell lining. Therefore, impaired responsiveness to EDV did not reflect a loss of endothelial cells due to desquamation.

One important feature of EC-LDL is its high content in amphiphiles, such as lysolipids, lipid peroxides and oxygenated sterols^{7,8}. Exchange of lipid between carriers in aqueous solution involves a rate-limiting movement of lipid through the aqueous phase¹². To determine whether effects of EC-LDL are related to transferable hydrophilic lipids, we have absorbed the lipoprotein with defatted albumin, a carrier with high avidity for lysophospholipids¹³. Albumin treatment reduced the elevated lysolecithin content of EC-LDL to that of N-LDL, and at the same time substantially attenuated its inhibitory effects on endothelium-dependent relaxation (Tables 1 and 2). To demonstrate that lipid in EC-LDL can account for its injurious effects on endothelial function, we extracted preparations of EC-LDL and N-LDL with chloroform-methanol (2:1 (V/V)) and dispersed the lipid extracts by sonication¹⁴. Lipid dispersions from EC-LDL suppressed the responsiveness to EDVs, whereas those from N-LDL were ineffective (Table 1). Furthermore, arteries incubated with lysolecithin (palmitoyl-lysolecithin) in submicellar concentrations ($\leq 10 \mu\text{M}$)¹⁵ produced changes in the reactivity to vasodilators closely resembling those obtained with EC-LDL or its extracts. Because lysolecithin acts as an endothelium-dependent vasodilator¹⁶, we wondered whether refractoriness to EDVs after EC-LDL or lysolecithin could represent a desensitization response. But incubations with EC-LDL ($n=7$) or lysolecithin ($n=7$) elicited only minor unsustained reductions in phenylephrine-induced constrictor tone, an effect unlikely to be the cause of profound and sustained unresponsiveness.

Because effects of amphiphiles on red blood cells have been well characterized, we have evaluated membrane-active effects of our lipoprotein and lipid preparations on red blood cells¹⁷. Incorporation of lysolecithin in red cell membranes produces a quantifiable transformation of the erythrocytes from their discoidal to their crenated (echinocytic) configuration¹⁷. EC-LDL or its extracted lipid acted as potent echinocytogenic agents, whereas N-LDL, its lipid, and albumin-treated EC-LDL were

TABLE 2 Lysolecithin in LDL preparations

	Lecithin (nmol per mg protein)	Lysolecithin (nmol per mg protein)
Native LDL	1,470 $\pm$ 110	20 $\pm$ 5
EC-LDL (oxidized LDL)	610 $\pm$ 90	590 $\pm$ 40
A-EC-LDL (albumin-treated EC-LDL)	540 $\pm$ 50	20 $\pm$ 6
Albumin, after treatment*	0.8 $\pm$ 0.2	3.6 $\pm$ 0.1

Lipid extracted from LDL preparations ($n=6$) and albumin ($n=6$) were analysed by TLC using silica-gel G plates developed with a solvent mixture containing chloroform-methanol-water (25:10:1 by volume)³. Lipid bands representing lecithin and lysolecithin were eluted, and phospholipid phosphorus in eluates was quantitated by the Bartlett procedure⁸. With the LDL to albumin protein ratio of 1/100 fractional recovery of lysolecithin in the albumin fraction is $(3.6 \times 100/590 - 20) = 0.63$. Recovery of albumin-containing infranant was incomplete to avoid contamination with supernatant. Values are means $\pm$ s.e.m. of six to seven independent experiments.

* Defatted albumin before treatment contained no detectable phospholipid.

virtually ineffective (Table 1). Lysolecithin (palmitoyl lysolecithin) in concentrations approaching those obtained in experiments with EC-LDL had a crenating potency comparable to that of modified lipoprotein. Therefore, for all lipoprotein and lipid preparations tested, there was an apparent concordance between membrane-active effects on endothelial cells and those on erythrocytes.

An action of EC-LDL and its lipid on the responsiveness to EDVs involving disparate endothelial surface receptors despite a partially preserved reactivity to the receptor-independent calcium ionophore, indicates that lipid transferred to endothelial cell membranes interferes with an intramembrane (noncytosolic) regulatory pathway shared by the different receptor mechanisms. Because the EDVs that we tested are regulated by G proteins, the lipid could perturb G protein-dependent transmembrane signalling. In this context, it is of interest that lysolecithin, but not lecithin, activates GTPase activity of the G protein-like Ras protein¹⁸.

Inhibitory effects of high concentrations of LDL (2,000 μg LDL protein ml^{-1}) on endothelium-dependent arterial relaxation have been reported by Andrews *et al.*¹⁹. These authors suggested that the effects are mediated by an apoprotein-receptor mechanism involving native (unoxidized) LDL. This conclusion, however, was not based on receptor pharmacological analyses with labelled apoB, and lysolipids in LDL were not measured.

In conclusion, our results indicate that lysolecithin, an amphiphile that accumulates in modified LDL and in atherosclerotic arterial walls²⁰, is capable of producing a defect in endothelium-dependent vasomotor regulation that mimics that observed in atherosclerotic arteries. □

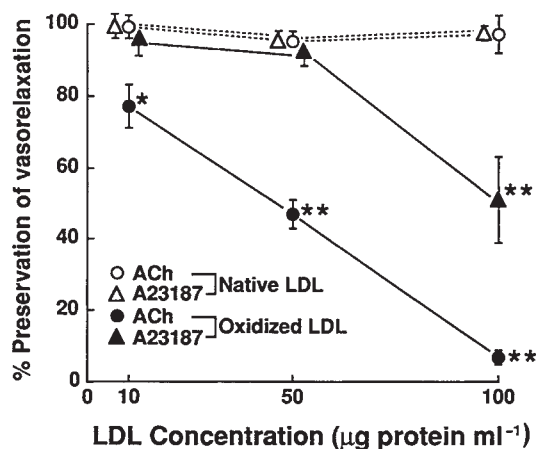


FIG. 2 Experiments with isolated arteries were performed as described in Fig. 1 and Table 1. Relaxations elicited after incubation with varying concentrations of N-LDL or EC-LDL are expressed as a percentage of the corresponding pre-incubation value. Drugs were used in concentrations yielding maximal relaxation. Data points represent means $\pm$ s.e.m. of six to nine independent experiments. For significant differences (*, $P < 0.02$; ** $P < 0.01$; unpaired t test), values were compared with corresponding N-LDL values.

Received 24 August 1989; accepted 10 January 1990.

- Habib, J. B. *et al.* *Circulation Res.* **58**, 305-309 (1986).
- Bossaller, C. *et al.* *J. clin. Invest.* **79**, 170-174 (1987).
- Yamamoto, H., Bossaller, C., Cartwright, J. & Henry, P. D. *J. clin. Invest.* **81**, 1752-1758 (1988).
- Furchgott, R. F. & Vanhoutte, P. M. *FASEB J.* **3**, 2007-2018 (1989).
- Yokoyama, M. & Henry, P. D. *Circulation Res.* **45**, 479-486 (1979).
- Henry, P. D. *Circulation* **72**, 456-459 (1985).
- Steinberg, D., Parthasarathy, S., Carew, T. E., Khoo, J. C. & Witztum, J. L. *New Engl. J. Med.* **320**, 915-924 (1989).
- Steinbrecher, U. P., Parthasarathy, S., Leake, D. S. & Witztum, J. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3883-3887 (1984).
- Henriksen, T., Mahoney, E. M. & Steinberg, D. *Arteriosclerosis* **3**, 149-159 (1983).
- Parthasarathy, S., Steinbrecher, R. P., Barnett, J., Witztum, J. L. & Steinberg, D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3000-3004 (1985).
- Yia-Herttuala, S. *et al.* *J. clin. Invest.* **84**, 1086-1095 (1989).
- Pownall, H. J., Hickson, D. L. & Smith, L. C. *Am. Chem. Soc.* **105**, 2440-2445 (1983).
- Mohandas, N., Wyatt, J., Mei, S. F., Rossi, M. E. & Shohet, S. B. *J. biol. Chem.* **257**, 6537-6543 (1982).
- Cooper, R. A., Arner, E. C., Wiley, J. S. & Shattil, S. J. *J. clin. Invest.* **55**, 115-126 (1975).
- Weltzien, H. U., Arnold, B. & Reuther, R. *Biochim. biophys. Acta* **466**, 411-421 (1977).
- Saito, T., Wolf, A., Mennon, N. K., Saeed, M. & Bing, R. J. *Proc. natn. Acad. Sci.* **85**, 8246-8250 (1988).
- Ferrell, J. E. Jr, Lee, F. K. & Huestis, W. H. *Biochemistry* **24**, 2849-2857 (1985).
- Tsai, S.-C. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 8310-8314 (1985).
- Andrews, H. E., Bruckdorfer, K. R., Dunn, R. C. & Jacobs, M. *Nature* **327**, 237-239 (1987).
- Vidaver, G. A., Ting, A. & Lee, J. W. *J. theor. Biol.* **115**, 27-41 (1985).

Self-tolerance to transgenic $\gamma\delta$ T cells by intrathymic inactivation

Marc Bonneville*, Isao Ishida, Shigeyoshi Itoharu, Sjeef Verbeek†, Anton Berns†, Osami Kanagawa‡, Werner Haas§ & Susumu Tonegawa

Howard Hughes Medical Institute at Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Division of Molecular Genetics of the Netherlands Cancer Institute, and Department of Biochemistry of the University of Amsterdam, Amsterdam, The Netherlands

‡ Eli Lilly Research Laboratories, La Jolla, California 92037, USA

§ Hoffman LaRoche AG, Central Research Unit, Basel, Switzerland

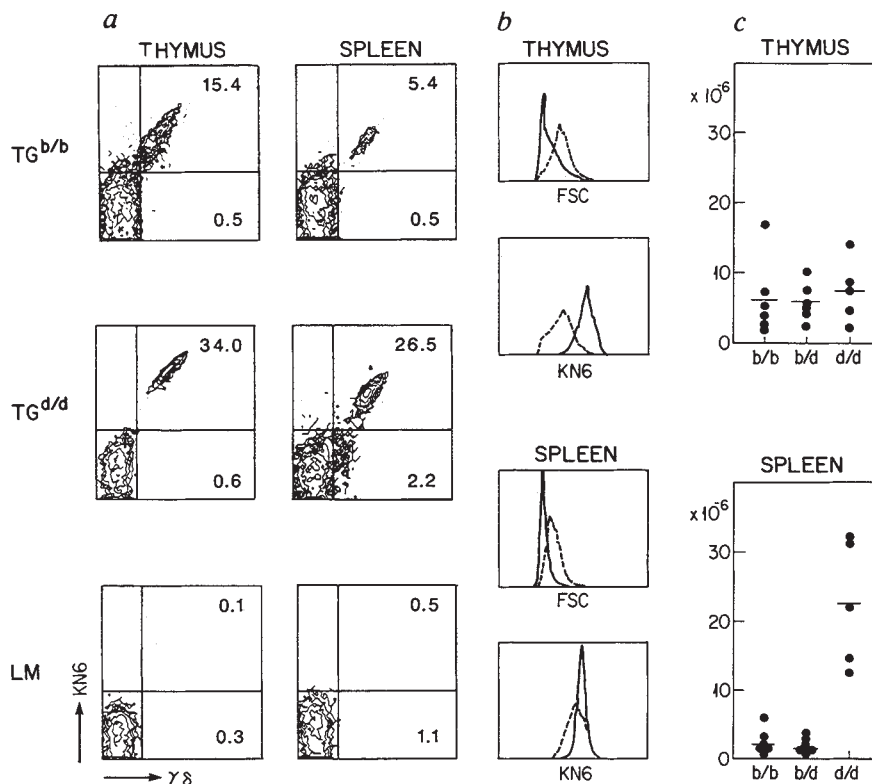
* Present address: Faculté de Médecine, INSERM U211.1, Rue G. Veil, 44035 Nantes Cedex, France

DURING their intrathymic differentiation, T lymphocytes expressing $\alpha\beta$ T-cell receptors (TCR) are negatively¹⁻³ and positively selected⁴⁻⁸. This selection contributes to the establishment of self-tolerance and ensures that mature CD4⁺ and CD8⁺ cell populations are restricted by the self major histocompatibility complex. Little is known, however, about $\gamma\delta$ T-cell development. To investigate whether selection operates in the establishment of the $\gamma\delta$ T-cell class, we have generated transgenic mice using γ - and δ -transgenes encoding a TCR that is specific for a product of a gene in the TL-region of the TL^b haplotype. Similar numbers of

thymocytes expressing the transgenic TCR were generated in mice of TL^b and TL^d haplotypes. But $\gamma\delta$ thymocytes from TL^b and TL^d transgenic mice differed in cell size, TCR density and in their capacity to respond to TL^b stimulator cells or interleukin-2 (IL-2). In contrast to $\gamma\delta$ T cells from TL^d transgenic mice, $\gamma\delta$ T cells from TL^b transgenic mice did not produce IL-2 and did not proliferate in response to TL^b stimulator cells, but they did proliferate in the presence of exogenous IL-2. These results indicate that functional inactivation of self-antigen-specific T cells could contribute to the establishment of self-tolerance to thymic determinants.

Several $\gamma\delta$ TCR-positive T hybridomas were obtained by fusing CD4⁺CD8⁺ thymocytes from adult C57BL/6 (H-2^b, TL^b) mice with the thymoma BW5147 (ref. 9). Using a hybridoma growth inhibition assay, we identified one hybridoma (KN6) that was specific for syngeneic cells¹⁰. It recognized a product of the TL^b haplotype, but not a product of the TL^d haplotype, that was expressed not only on peritoneal and spleen cells, but also in the thymus¹⁰. We were therefore interested in whether the KN6 hybridoma originated from an immature $\gamma\delta$ T cell destined to be eliminated in the thymus, but was rescued by the fusion. To investigate this possibility, we generated transgenic mice by introducing into the germ line the rearranged V4J1C1 γ - and V5DJ1C8 δ -chain genes encoding the KN6 TCR¹⁰⁻¹². We backcrossed CBA/J $\times$ C57BL (TL^{k/b}) transgenic mice several times to C57BL/6 (TL^b) or BALB/c (TL^d) mice, and analysed homozygous TL^{b/b} and TL^{d/d} mice, as well as heterozygous TL^{b/d} transgenic mice. We studied $\gamma\delta$ TCR expression by immunofluorescence using monoclonal antibodies that react

FIG. 1 Transgene-encoded $\gamma\delta$ TCR are expressed on the surface of $\gamma\delta$ T cells from KN6 transgenic mice of TL^{b/b} and TL^{d/d} haplotypes. **a**, Thymus and spleen cell suspensions were stained with fluorescein isothiocyanate (FITC)-conjugated 3A10 monoclonal antibody (anti-pan $\gamma\delta$)¹³ and biotinylated 8D6 monoclonal antibody (anti-V γ 4V δ 5-encoded TCR)¹³, followed by streptavidin coupled to phycoerythrin. The dot-plot histograms shown (3A10, horizontal axis; 8D6, vertical axis) were obtained after flow cytometric analysis of thymocytes and splenocytes from KN6 transgenic mice of TL^b (TG^{b/b}) and TL^d (TG^{d/d}) haplotypes and from nontransgenic littermate is (LM) of TL^d haplotype mice. Fluorescence intensity is expressed in log₁₀ units. The percentage of cells in right quadrants are indicated. **b**, TCR density and size of $\gamma\delta$ T cells from KN6 transgenic mice. Thymus and spleen cell suspensions were stained with 3A10 and 8D6 monoclonal antibodies according to **a**. The forward scatter (FSC) and the 8D6 fluorescence histograms (TL^{b/b} transgenic, dotted line; TL^{d/d} transgenic, solid line) were obtained from gated $\gamma\delta$ T cells. **c**, Total number of KN6 $\gamma\delta$ TCR-positive cells in thymus and spleen from TL^{b/b}, TL^{b/d} and TL^{d/d} KN6 transgenic mice. Single-cell suspensions of thymic and splenic cells from KN6 transgenic mice (4–10 week-old) were stained with biotinylated monoclonal antibody 5C10 (KN6 clonotypic) followed by streptavidin coupled to phycoerythrin, and the number of KN6 $\gamma\delta$ TCR-positive cells was calculated from the percentage of 5C10⁺ cells and the total cell number. The obtained mean values (horizontal bars) and s.d.s. were the following (in 10⁶ cells): thymus (TL^{b/b}, 6.0 $\pm$ 5.1; TL^{b/d}, 5.9 $\pm$ 2.4; TL^{d/d}, 7.5 $\pm$ 4.4); spleen (TL^{b/b}, 2.2 $\pm$ 1.9; TL^{b/d}, 1.9 $\pm$ 1.0; TL^{d/d}, 22.4 $\pm$ 9.4). **METHODS.** KN6 $\gamma\delta$ transgenic mice were generated after injection into fertilized eggs of purified V4J1C1 γ and V5DJ1C8 δ cosmid DNA from the KN6 hybridoma as previously described^{11,12}. The CBA/J $\times$ C57BL transgenic founders obtained were crossed to C57BL/6J and BALB/CByJ mice (Jackson Laboratories) and their progeny backcrossed to their respective strains. Major histocompatibility complex typing was performed on cultured peripheral blood lymphocytes using 28-13-3S (anti-H-2K^b), SF1.1.1 (anti-H-



2K^d) and 16.1.11N (anti-H-2K^k) monoclonal antibodies (ATCC No. HB41, HB159 and HB16, Rockville, Maryland). Spleen and thymus cell suspensions from 6-week-old transgenic and nontransgenic mice were prepared according to standard procedures and stained with conjugated 3A10 and 8D6 monoclonal antibodies or with unconjugated 5C10 (anti-KN6 TCR clonotype)¹³ monoclonal antibody as described¹³. Flow cytometric analysis was performed on a FACSCAN (Becton Dickinson) using Consort C30 and LYSYS programs. The percentages of 8D6⁺ and 5C10⁺ cells were virtually identical (data not shown).

with all $\gamma\delta$ TCR (3A10)¹³, with most or all $V\gamma 4V\delta 5$ -encoded TCR (8D6)¹³, or with the KN6 $\gamma\delta$ TCR only (5C10)¹³. Virtually all thymic and splenic $\gamma\delta$ T cells from TL^{d/d} or TL^{b/b} transgenic mice expressed the transgene-encoded TCR (Fig. 1a), due to blockade of endogenous γ - and δ -gene rearrangement¹⁰⁻¹². Anti-CD4 and anti-CD8 monoclonal antibodies did not stain any KN6 TCR-positive cells (data not shown). The total number of peripheral and thymic $\alpha\beta$ T cells varied in different litters, and will not be further considered here. By contrast, the numbers of KN6 TCR-positive thymocytes was rather constant and not significantly different between transgenic mice of the TL^{b/b} and TL^{d/d} haplotypes (Fig. 1c). But the numbers of splenic KN6 TCR-positive cells was on average 10 times lower in TL^{b/b} or TL^{b/d} transgenic mice than in TL^{d/d} transgenic mice. Also $\gamma\delta$ T cells in TL^{b/b} or TL^{b/d} transgenic mice were larger than those in TL^{d/d} transgenic mice, and expressed surface KN6 TCR at a lower density. This difference was more pronounced in thymocytes than in spleen cells (Fig. 1b).

To assess functional properties of $\gamma\delta$ T cells in the various TCR transgenic mice, we purified these cells from thymus and spleen by treatment with anti-CD4 and anti-CD8 monoclonal antibodies and complement. The resulting cell population (<1% $\alpha\beta$ T cells, >90% $\gamma\delta$ T cells) was cultured either in medium alone or in medium with TL^b stimulator cells or IL-2, or both. No spontaneous proliferation was observed in medium alone (Table 1 and Fig. 2a). IL-2 receptor α -chain expression (Fig. 3) and proliferation (Fig. 2) were induced in thymocytes from TL^{b/b} or TL^{b/d} transgenic mice by IL-2 but not by TL^b stimulator cells. TL^b stimulator cells also failed to induce IL-2 or IL-4 production, or both, by TL^{b/b} thymocytes (Table 1). Opposite results were obtained with thymocytes from TL^{d/d} transgenic mice: IL-2 or IL-4 production, or both, proliferation and IL-2

receptor α -chain expression were induced by TL^b stimulator cells but not by IL-2 (Table 1, and Figs 2 and 3). In addition, although thymocytes from both TL^{b/b} and TL^{d/d} transgenic mice proliferated and expressed IL-2 receptor α -chain when exposed to IL-2 and TL^b stimulator cells (Figs. 2d and 3d, h), the response of the latter was much stronger. Essentially the same results were obtained with spleen cells, except splenic $\gamma\delta$ cells from TL^{b/b} transgenic mice did not respond to IL-2 alone (data not shown). These results indicate that the anergized splenic $\gamma\delta$ T cells are resting.

The results described above show that a self-antigen that is expressed in the thymus does not necessarily delete $\gamma\delta$ T cells expressing TCR for that antigen; we did not see any reduction in the number of KN6 TCR-positive thymocytes in TL^b transgenic mice relative to that in TL^d transgenic mice. This finding differs from the results of studies with $\alpha\beta$ TCR transgenic mice which demonstrates clearly the deletion of self-antigen-reactive cells^{3,6}. In all these previous studies, the TCR genes that were used for the construction of the transgenic mice were obtained from *in vitro*-stimulated T-cell clones, whereas the KN6 TCR-encoding genes expressed in the transgenic mice that we used were obtained from a $\gamma\delta$ TCR-positive thymocyte. It is possible that the affinity of the transgenic TCR that we were using for a product of the TL^b haplotype was below the threshold required for deletion of immature cells as well as for activation of mature cells. This possibility seems unlikely, because the affinity of the KN6 TCR is sufficient to mediate a proliferative response of mature $\gamma\delta$ T cells from TL^d mice to TL^b stimulator cells. It is more likely that it is the nature of the cell presenting an antigen in the thymus that determines whether T cells reactive with that antigen are deleted or inactivated. It is possible that a minor subset of KN6 TCR-positive thymocytes is in fact deleted in

TABLE 1 Functional properties of $\gamma\delta$ T cells from TL^{b/b} KN6 transgenic mice

Transgenic responder cells (CD4 ⁺ CD8 ⁻)	TL ^b stimulator cells	³ H-labelled thymidine uptake (c.p.m.)	IL-2/4 production (c.p.m.)
TL ^{d/d} thymocytes	-	400 ± 100	1,600 ± 500
	+	50,500 ± 4,700	31,100 ± 1,800
TL ^{b/b} thymocytes	-	700 ± 300	3,400 ± 800
	+	1,500 ± 400	1,200 ± 300
TL ^{d/d} splenocytes	-	1,600 ± 400	3,300 ± 1,100
	+	55,200 ± 2,600	27,300 ± 500
TL ^{b/b} splenocytes	-	1,000 ± 300	1,100 ± 300
	+	2,100 ± 600	1,400 ± 400

When exposed to specific antigen-bearing cells, $\gamma\delta$ T cells from TL^{b/b} KN6 transgenic mice are unable to proliferate or secrete lymphokines. CD4⁺CD8⁻ thymocytes and splenocytes from TL^{d/d} and TL^{b/b} KN6 transgenic mice were cultured with or without irradiated peritoneal cells of C57BL/6J origin (TL^b). Proliferative activity of responder cells, assessed by ³H-labelled thymidine uptake, was assayed at day 2 after the initiation of the culture. Culture supernatants were collected at day 1 and tested for their ability to trigger proliferation of IL-2/4-dependent HT-2 cells. The experiment was performed three times, obtaining similar results. A representative result from one of the three experiments is shown. All measurements were made in triplicate and the data are expressed as mean ± s.d. Spleen and thymus cell suspensions from 6-week-old TL^{b/b} and TL^{d/d} KN6 transgenic mice were incubated for 1 h at 37 °C with anti-CD4 (RL172.4) and anti-CD8 (3.155.D14) monoclonal antibodies and complement, and live cells were recovered by centrifugation over Ficoll-hypaque. Double-negative responder cells (5 × 10³) were cultured in 96-well plates in 200 µl complete culture medium (RPMI 1640, 10% FCS, 2 mM L-glutamine) with or without 2.5 × 10⁴ irradiated (1,500 rad) peritoneal cells from C57BL/6J mice. Culture supernatant (100 µl) was collected at day 1. At day 2, cultures were pulsed for 6 h with [³H-methyl] thymidine (1 µCi per well), collected and counted in a scintillation counter. To determine IL-2/4 production, dilutions of day-1 culture supernatants were added to culture of IL-2/4-dependent HT-2 cells. [³H-methyl] thymidine incorporation by these cells was determined on day 2. The results shown are the mean c.p.m. ± s.d. of triplicates that were obtained with 1/4 final dilutions.

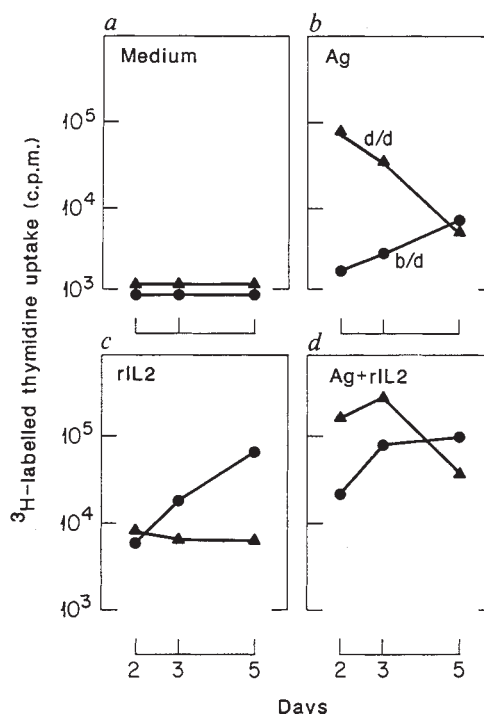


FIG. 2 Addition of recombinant IL-2 (rIL-2) partly restores the proliferation of $\gamma\delta$ T cells from TL^{b/b} KN6 transgenic mice to TL^b stimulator cells. CD4⁺CD8⁻ thymocytes (TL^{b/b} transgenic, ●; TL^{d/d} transgenic, ▲) were cultured without (a, c) or with (b, d) irradiated TL^b peritoneal cells (Ag), in the absence (a and b) or the presence (c and d) of rIL-2. Their proliferative activity, assessed by ³H-labelled thymidine uptake, was assayed at day 2, 3 and 5 after the initiation of the culture. METHODS. CD4CD8 double-negative responder cells and irradiated TL^b peritoneal cells were prepared and cultured as described in Table 1. IL-2 was added to a final concentration of 100 U ml⁻¹. All cultures were pulsed with 1 µCi of [³H-methyl]thymidine 6 h before collection.

the thymus of TL^b mice. But it is clear that most of these cells are not eliminated even though they interact with products of the TL^b haplotype in the thymus. As a result of this interaction, KN6 TCR-positive cells increase in size, down-modulate their TCR and probably express IL-2-receptor β -chains, which would confer on them the ability to respond to IL-2 alone¹⁴. This effect of the TL^b product does not reflect an essential positive selection step, because KN6 TCR-positive cells do survive in, and are exported from, the thymus of TL^d transgenic mice. Whereas KN6 TCR-positive cells from TL^d thymi respond to TL^b stimulator cells with IL-2 production and proliferation, KN6 TCR-positive cells from TL^b thymi lose the capacity to produce IL-2 and, as a result, respond to TL^b stimulator cells only in the presence of exogenous IL-2. Therefore the response of KN6 TCR-positive cells from TL^b mice to TL^b stimulator cells is dependent on helper cells that supply IL-2. The lack of such helper cells in TL^b mice could explain the lack of destructive

autoimmune responses in TL^b transgenic mice and the lower number of transgenic TCR-positive cells in the periphery of TL^b mice versus TL^d mice. The unresponsive state of the TG-positive cells that is generated in the thymus of the transgenic mice used in the study described here is reminiscent of the state of clonal energy that can be induced in peripheral $\alpha\beta$ T cells¹⁵⁻¹⁸, as well as in the thymocytes of P $\rightarrow$ F₁ chimaeras¹⁹.

Our study shows that certain T cells may become dependent on helper cells if they encounter self-antigen in the thymus as well as in the periphery. Like self antigen-specific B cells, self antigen-specific T cells are not necessarily harmful to the host. In fact they could have beneficial functions provided that helper cells are absent, which would facilitate their proliferation and differentiation to potentially harmful effector cells in response to host components. □

Received 24 November 1989; accepted 10 January 1990.

1. Lederberg, J. *Science* **129**, 1649-1653 (1959).
2. Kappler, J., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
3. Kiselow, P., Bluthman, H., Staerz, U., Steinmetz, M. & von Boehmer, H. *Nature* **333**, 742-748 (1988).
4. Sprent, J., Lo, D., Gao, E. H. & Ron, Y. *Immunol. Rev.* **101**, 173-190 (1988).
5. Teh, H. S. *et al. Nature* **335**, 229-233 (1988).
6. Sha, W. C. *et al. Nature* **335**, 271-274 (1988).
7. Blackman, M., Marrack, P. & Kappler, J. *Science* **244**, 214-217 (1989).
8. Berg, L. J. *et al. Cell* **58**, 1035-1046 (1989).
9. Takagaki, Y., Nakanishi, N., Ishida, I., Kanagawa, O. & Tonegawa, S. *J. Immunol.* **142**, 2112-2121 (1989).
10. Bonneville, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 5928-5932 (1989).
11. Ishida, I. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
12. Tonegawa, S. *et al. Cold Spring Harb. Symp. quant. Biol.* (in the press).
13. Itohara, S., Nakanishi, N., Kanagawa, O., Kubo, R. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5094-5098 (1989).
14. Toribio, M. L., Gutierrez-Ramos, J. G., Pezzi, L., Marcos, M. A. R. & Martinez-A. C. *Nature* **342**, 82-85 (1989).
15. Schwartz, R. H. *Cell* **57**, 1073-1081 (1989).
16. Rammensee, H.-G., Kroschewski, R. & Frangoulis, B. *Nature* **339**, 541-544 (1989).
17. Win, D. *et al. J. exp. Med.* **169**, 779-794 (1989).
18. Lo, D. *et al. J. exp. Med.* **170**, 87-104 (1989).
19. Ramsdell, F., Lantz, T. & Fowlkes, B. J. *Science* **246**, 1038-1041 (1989).

ACKNOWLEDGEMENTS. We thank Edvins Krecko, Patricia Halpan and Mike Surman for their technical assistance, M. Nabholz, D. Sachs, K. Ozato and J. Stall for monoclonal antibodies, and A. Coutinho for discussions. We also thank Anneloes Beenders and Nel Bosnie for raising and analysing the transgenic mice. This work was supported by the Howard Hughes Medical Institute, Ajinomoto Ltd, the American Cancer Society, and the NIH (S.T.); The Dutch Cancer Society (K.W.F.), The Netherlands Cancer Foundation (S.V. and A.B.); and the NIH CORE (P. Sharp). I.I. was supported in part by Toyobo Biotechnology Foundation. M.B. is supported in part by the Association pour la Recherche contre le Cancer and Ligue Nationale contre le Cancer.

A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus

Sven Pettersson*, Graham P. Cook, Marianne Brüggemann*, Gareth T. Williams & Michael S. Neuberger

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

THE expression of immunoglobulin heavy-chain (IgH) genes is generally thought to be regulated by the combination of the V_H promoter with the enhancer element which is located in the JH-CH intron¹⁻⁴. This is probably an oversimplification: there are cell lines that transcribe IgH genes despite the deletion of the intron-enhancer⁵⁻⁸. These findings could imply that other enhancer element(s) exist in the IgH locus⁹⁻¹¹. Here we show that a strong B-cell-specific enhancer is indeed located at the 3'-end of the rat IgH locus, 25 kilobases downstream of C α . This enhancer should be retained downstream of all rearranged IgH genes, regardless of the V_H or CH segment used. Taken together with analogous findings for the mouse κ locus¹², the results prompt a re-evaluation

* Present addresses: Centre for Biotechnology, Blickagängen 6, 14152 Huddinge, Sweden (S.P.); IAPGR, Babraham, Cambridge CB2 4AT, UK (M.B.).

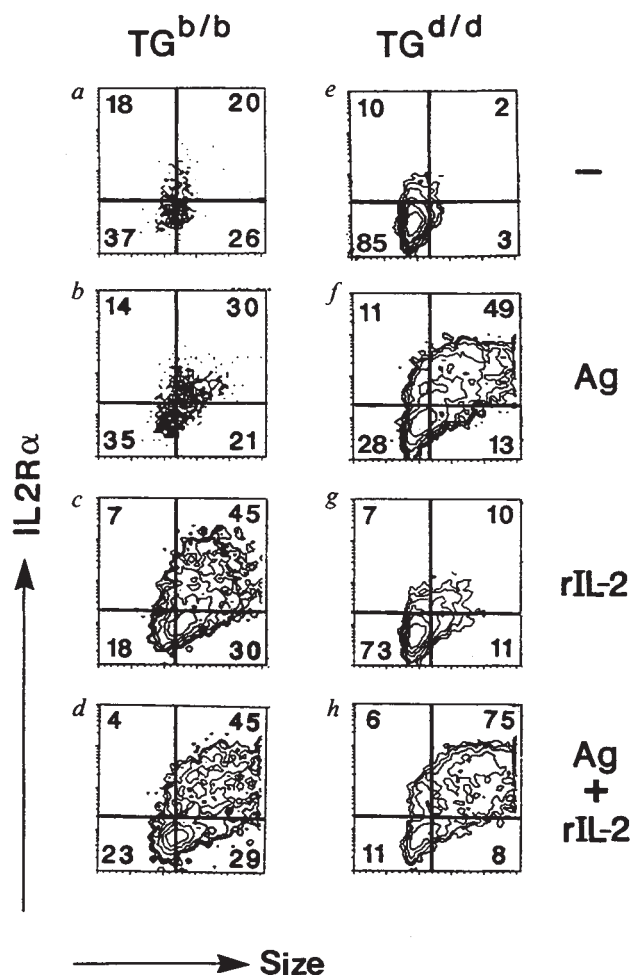


FIG. 3 Blastogenesis and IL-2 receptor α -chain (IL-2R α) expression of cultured DN thymocytes from KN6 TG mice. DN thymocytes from transgenic mice of TL^b (TG^{b/b}) or TL^d (TG^{d/d}) haplotypes were cultured in culture medium alone (a, e), with irradiated TL^b peritoneal cells (b, f), with rIL-2 alone (c, g), or with rIL-2 and TL^b stimulator cells (d, h). After 2 days of culture, cells were stained with 8D6 (anti-V γ 4V δ 5) and PC61 (anti-IL-2R α) monoclonal antibodies. Shown are the dot-plot histograms (IL-2R α fluorescence intensity (log₁₀ scale) on the vertical axis, and forward scatter (FSC; linear scale) on the horizontal axis of gated 8D6⁺ cells, and the percentage of cells in each quadrant.

METHODS. Cultured cells were incubated with PC61 monoclonal antibody (ATCC No. T1B222) followed by FITC-coupled purified goat anti-rat IgG antiserum (CALTAG, San Francisco). After a third incubation in 10% normal rat serum, cells were stained with biotin-coupled 8D6 monoclonal antibody followed by streptavidin-phycoerythrin. Flow cytometry was performed according to procedures described in Fig. 1 legend.

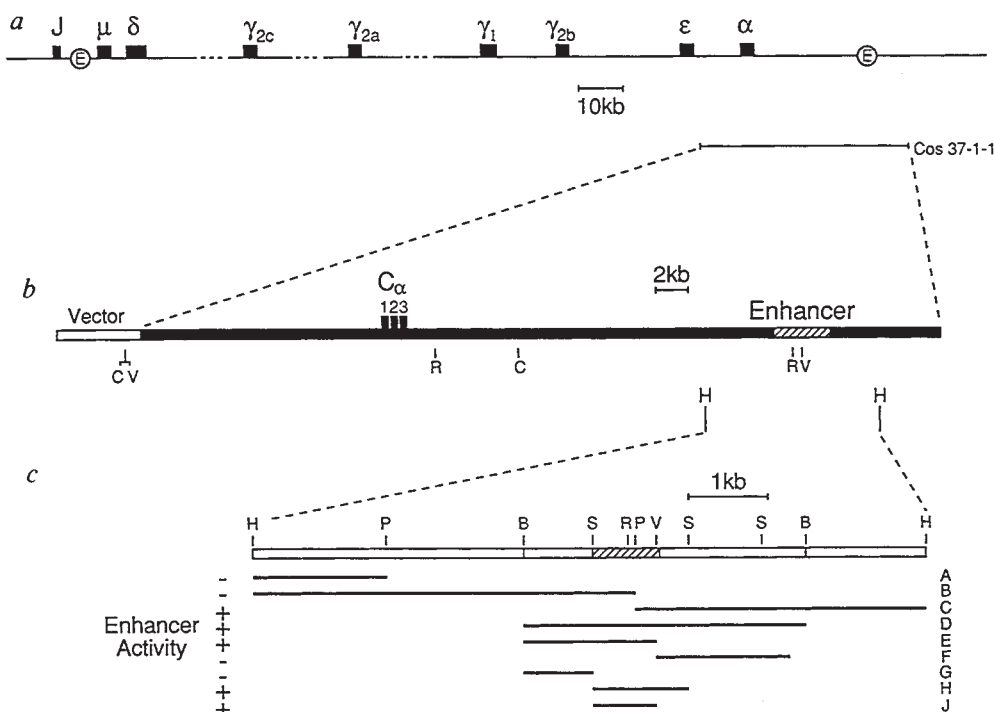


FIG. 1 *a*, The rat IgH locus. (E) denotes the intron and 3' enhancers. The interrupted line linking *Cy2c* and *Cy2a* indicates that the genes have not been physically linked; the gene order derives from analysis of gene deletions in hybridomas¹⁴; the rest of the rat IgH genes have been linked on overlapping cosmids¹³. *b*, Map of cosmid 37-1-1. *c*, Map of the *HindIII* enhancer fragment and the enhancer activity of subfragments. Assays were performed as described in the legend to Fig. 2. Restriction enzyme cleavage sites: B, *BglII*; C, *Clal*; H, *HindIII*; P, *PstI*; R, *EcoRI*; S, *StuI*; Sm, *SmaI* and V, *EcoRV*. The single *PstI* site indicated in the shaded enhancer fragment is a doublet.

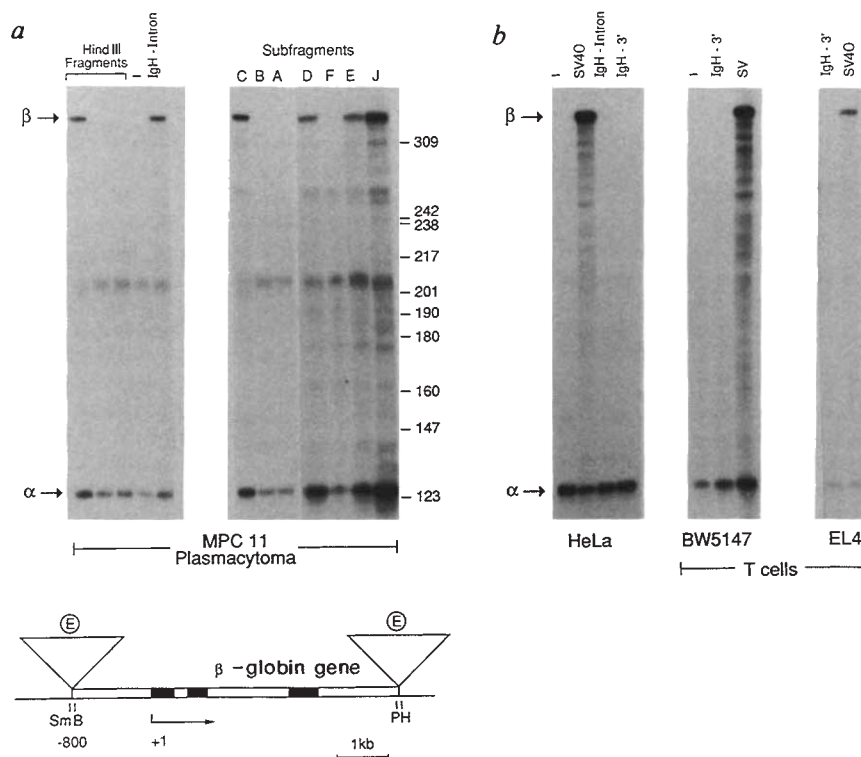
of the mechanism of regulation of immunoglobulin gene transcription. Furthermore, unlike the intron-enhancer, the IgH 3' enhancer would become linked to a *c-myc* that rearranges into an IgH switch region. The IgH 3' enhancer could therefore play a part in the activation of the translocated *c-myc* genes in rat immunocytomas, mouse plasmacytomas and Burkitt lymphomas.

In screening for an additional enhancer in the IgH locus, we focused our attention on sequences 3' of *Cα*, because this region is likely to be maintained in all rearranged IgH genes regardless of the particular VH or CH gene used. We used cosmids that covered >30 kilobases (kb) of the 3' end of the rat IgH locus. The locus of rat (Fig. 1*a*) has striking homology to that of the

mouse both in overall organization and in sequence. For example, the rat intron-enhancer and many of the individual CH genes show >90% homology to their mouse counterparts^{13,14}.

To test for enhancer activity, we exploited the fact that the human β -globin gene in plasmid p β 800 (Fig. 2*a*) is only a weak transcription unit in transfected cells unless it is provided with an exogenous enhancer element. We scanned cosmid 37-1-1 for enhancer elements because this cosmid spans *Cα* and the region 3' of it (Fig. 1*a*). Random *HindIII* fragments of Cos 37-1-1 were inserted downstream of the β -globin gene in plasmid p β 800. The DNA from the resulting constructs was transfected into the

FIG. 2 *a*, Location of the rat IgH 3' enhancer. Ribonuclease protection assay measurements of human β - and α -globin mRNAs in transfected mouse MPC11 plasmacytoma cells. Cells were transfected with derivatives of plasmids p β 800 (see diagram) or p β 128, along with a reference plasmid that included the human α_2 -globin gene. Bands corresponding to correctly initiated β - and α -globin transcripts are indicated. Plasmid p β 800 (gift from K. Weston) consists of the human β -globin gene with 800 nucleotides (nt) of 5' flanking sequence cloned between the *HincII* and *PstI* sites of plasmid pUC12; plasmid p β 128 includes 128 nt of 5' flanking sequence. The derivatives of plasmid p β 800 include *HindIII* (left panel) or *PstI* fragments (fragments A, B and C) cloned downstream of β -globin gene, whereas the other fragments were cloned upstream. The mouse IgH-intron enhancer (IgH-Intron; 1-kb *XbaI* fragment) served as a positive control. *b*, Cell-type specificity of the IgH 3' enhancer (IgH-3'; *StuI*-*EcoRV* fragment). Parallel assays were performed in HeLa cells and in mouse thymoma cell lines EL4 and BW5147. The SV40 enhancer served as a positive control. Transfection of MPC11 and HeLa cells by calcium phosphate co-precipitation, transfection of T-cell lymphomas by use of DEAE-dextran, and measurements of RNA levels, were carried out as previously described^{4,12}.



mouse plasmacytoma cell line MPC11 by calcium phosphate co-precipitation along with a human $\alpha 2$ -globin plasmid that served as an internal reference. The amount of β - and α -globin messenger RNAs produced by the transfected cells was then measured in ribonuclease protection assays. One of the *Hind*III fragments caused a considerable stimulation of β -globin transcription and gave rise to mRNA levels comparable to those achieved with the IgH intron-enhancer (Fig. 2a). Mapping of this *Hind*III fragment revealed that it contained sites for *Eco*RI and *Eco*RV, which allowed us to identify its position relative to the *C α* gene (Fig. 1b). Further fragmentation of the cosmid and of clones derived from it demonstrated that enhancer activity was in a 0.7-kb *Stu*I-*Eco*RV fragment located 25 kb downstream of the *C α* exons (Fig. 1c). This IgH 3' enhancer has the classic properties of a transcriptional enhancer element in that it is active in both orientations and functions both upstream and downstream of the test gene. (The *Hind*III fragments and *Pst*I fragments A, B and C were assayed downstream of the β -globin gene, whereas fragments D-J were assayed upstream. Similarly, fragments H and J were assayed in opposite orientations to each other). To test whether the 3' enhancer possessed ubiquitous or lymphoid-specific activity, we introduced the plasmid constructs into fibroblasts and T-cell lymphomas. Although activity was easily detected in plasmacytoma, no activity was detected in either HeLa cells or in the mouse T-cell lymphomas BW5147 and EL4 (Fig. 2b). Therefore, like the IgH-intron enhancer⁴, the 3' enhancer is active in plasmacytoma, but not active in T cells or nonlymphoid cells in these transient transfection assays.

The sequence of the *StuI-EcoRV* fragment which encompasses enhancer activity was determined and is shown in Fig. 3. It includes a stretch of GA repeats preceded by oligo(G-T).

StuI
AGGCCTGTGTCCATGTGGCCACAGCTAGCTCCCAGCCTGCCAAGCCTGTCTGTGGAGGGA
20 40 60

TTGCGGAGTGGGGCTGGGCTGAGGCAGGCCAGGATTTTCCAGTGGAGCAGCCAGTGACT
80 100 120

GGGCTGAGGGTCACTGCCACCATTCCCGTGATTCTGGGGTAGACCCAGGACCTGGAGCAC
140 160 180

ACACAGCCTTCTCTGTCTCTCCCACTTCCCTGGGGTGTTTGGACCTTGCCTATCTCTGTC
200 220 240

ATTCCCTCTAGGGTCAGATTCCAGCAGTGTGCATAAATGGGAAAACACTGAATTCATG
260 280 300

TCACCGTTAGGGGTCTAGGAACCCCGAGCCCCATCCCCAAAGCTGCTCAACCTGGCCAGG
320 340 360

GTGGGGTGAACCTGCAGCTGCAGGTGGGGAATCATTGATGCTATTTTGGCTCAGAGAA
380 400 420

TGAGAAACATGTTTTCTCATGGAAACAAATCATTATTCATGTGCCCCGAGGGCCAAC
440 460 480

CCTGCACAGAGTCAGTCACTCGGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAGAA
500 520 540

GAGAGAGAGAGAGAGAGAGAGAGAGAGACAGACAGAGACAGACAGACAGACAGAGA
560 580 600

GACAGAGAGACAGAGAGAGACAGAGAGACACACACAGAGACAGAGACAGACAGAGA
620 640 660

GACAGAGACAGAGAGACACACACAGAGACAGAGACAGAGAGAGATGTTGTGA
680 700 720

GCAAGCCCTGTTTCAGGTACTGATATC
740

EcoRI
EcoRV

Octanucleotide		HOMOLOGIES	
IgH-3'	456-ATTTCGAT	IgH-3' (invert)	386-CACCTGCAGC
Consensus	ATTTCGAT	IgH-intron	CACCTGCAGC
μE1		AP-1	
IgH-3'	346-GTCAACCTGGCC	IgH-3'	492-TGAGTCA
IgH-intron	GTCAAGATGGCC	Consensus	TGAGTCA
AP-2		AP-4	
IgH-3' (invert)	214-CCCCAGGG	IgH-3' (invert)	345-CAGCTTTGG
Consensus	CCCCAGGC	Consensus	CAGCTGTGG
IgH-3'	323-CCCCAGCC		

FIG. 3 Sequence of the rat IgH 3' enhancer (IgH-3'). The sequence of both strands was determined by the dideoxy-chain-termination method. Homologies or identities to regions of the mouse IgH-intron enhancer (IgH-intron) are boxed; those to AP-1, AP-2 or AP-4-binding sites²⁵ are underlined.

Comparison of the IgH 3' enhancer with the intron of the gene for IgH and simian virus 40 (SV40) enhancers revealed homologies that span several notable motifs. In particular, a perfect copy of the octanucleotide (ATTTCAT) is present at position 456. The octanucleotide element is found in the IgH-intron enhancer, is an essential component of VH promoters⁴ and is sufficient in appropriate assays to confer lymphoid-specific gene expression¹⁵⁻¹⁷. There is also a good match to another region of the mouse intron-enhancer that is implicated in lymphoid-specific transcriptional activation, that is the region surrounding the μ E2 and μ E5 motifs^{18,19}, as well as significant homology to a region surrounding the NF- μ E1 binding site. As regards SV40 comparisons, there is a match to the consensus binding site for transcription factor AP-1, and homologies to the AP-2- and AP-4-binding sites. Analysis of cosmid subfragments A, B and C (Fig. 1c) indicates that the most important motifs are likely to lie across, or 3' of, the *Pst*I site at position 378; this would include the octanucleotide, μ E5 and AP-1 motifs. But functional assays are clearly necessary to evaluate the significance of the homologies.

The rat IgH-intron enhancer shows B cell-specific enhancer activity and >90% sequence homology to its mouse counterpart (ref. 13, and our unpublished observations). Therefore, the rat IgH locus, like the mouse κ light-chain locus¹², contains both a J-C intron-enhancer and a 3' enhancer. We do not believe that the presence of several enhancers is peculiar to the rat IgH locus; preliminary studies indicate that a 3' IgH enhancer is also present in the mouse IgH locus. The presence of several enhancers in the immunoglobulin gene loci could simply reflect redundancy. From transgenic mouse studies²⁰, however, it seems that the combination of VH promoter with IgH intron-enhancer is not sufficient to give properly regulated IgH expression *in vivo*. In fact, in the context of the accessibility model of immunoglobulin gene rearrangement²¹, it is possible that the intron-enhancer is primarily concerned with the regulation of VH-D-JH joining, and that the 3' enhancer is implicated in transcriptional regulation at later stages of differentiation.

The existence of an enhancer at the 3' end of the IgH locus could explain the transcriptional activation of the translocated *c-myc* genes characteristic of several lymphoid malignancies. In

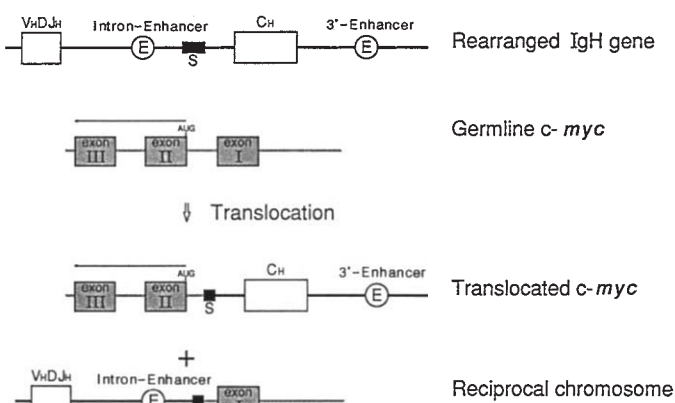


FIG. 4 The structure of a typical *c-myc*/IgH locus translocation. The precise structure of the translocation depends on the tumour, with the favoured incoming CH gene being *C ϵ* in rat immunocytoma, *C α* or *C γ* in mouse plasmacytoma, and *C μ* or *C γ* in Burkitt lymphoma. Similarly, the *c-myc* breakpoint can either be upstream of *c-myc* or it can be within *c-myc*, separating the noncoding exon (exon 1) from the coding region (the AUG initiator codon is in exon 2). In most cases, the breakpoint in the IgH locus is within a switch (S) region. The translocation therefore separates the IgH-intron enhancer from the coding region of *c-myc*, because they end up on reciprocal products of the translocation. But the translocated *c-myc* will become linked to an IgH 3' enhancer, which could therefore play a part in *c-myc* transcriptional activation.

many human Burkitt lymphomas, mouse plasmacytomas and rat immunocytomas, *c-myc* is translocated into the IgH locus^{22,23}. But the location of the breakpoints on the IgH locus-bearing chromosome indicates that the IgH-intron-enhancer is unlikely to be involved in the activation of the translocated *c-myc* allele, because the protein-coding region of *c-myc* and the IgH enhancer end up on reciprocal products of the translocation²⁴. As shown in Fig. 4, however, the translocated *c-myc* allele would typically be linked to the IgH 3' enhancer, and this enhancer could therefore be implicated in the origin or maintenance of B-cell neoplasias. □

Received 28 November 1989; accepted 29 January 1990.

1. Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729–740 (1983).
2. Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33**, 717–728 (1983).
3. Neuberger, M. S. *EMBO J.* **2**, 1373–1378 (1983).
4. Mason, J. O., Williams, G. T. & Neuberger, M. S. *Cell* **41**, 479–487 (1985).
5. Klein, S., Sablitzky, F. & Radbruch, A. *EMBO J.* **3**, 2473–2476 (1984).
6. Wabl, M. & Burrows, P. D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2452–2455 (1984).
7. Aguilera, R. J., Hope, T. J. & Sakano, H. *EMBO J.* **4**, 3689–3693 (1985).
8. Eckhardt, L. A. & Birshstein, B. K. *Molec. Cell Biol.* **5**, 856–868 (1985).
9. Klein, S., Gerster, T., Picard, D., Radbruch, A. & Schaffner, W. *Nucleic Acids Res.* **13**, 8901–8912 (1985).
10. Grosschedl, R. & Marx, M. *Cell* **55**, 646–654 (1988).
11. Schaffner, W. *Nature* **336**, 427–428 (1988).
12. Meyer, K. B. & Neuberger, M. S. *EMBO J.* **8**, 1959–1964 (1989).
13. Brüggemann, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6075–6079 (1986).
14. Brüggemann, M. *Gene* **74**, 473–482 (1988).
15. Gerster, T., Matthias, P., Thali, M., Jiricny, J. & Schaffner, W. *EMBO J.* **6**, 1323–1330 (1987).
16. Dreyfus, M., Doyen, N. & Rougeon, F. *EMBO J.* **6**, 1685–1690 (1987).
17. Wirth, T., Staudt, L. & Baltimore, D. *Nature* **329**, 174–178.
18. Ephrussi, A., Church, G. M., Tonegawa, S. & Gilbert, W. *Science* **227**, 134–140 (1985).
19. Murre, C., McCaw, P. S. & Baltimore, D. *Cell* **56**, 777–783 (1989).
20. Pettersson, S., Sharpe, M. J., Gilmore, D. R., Surani, M. A. & Neuberger, M. S. *Int. Immun.* **1**, 509–516 (1989).
21. Alt, F. W., Blackwell, T. K. & Yancopoulos, G. D. *Science* **238**, 1079–1087.
22. Pear, W. S. et al. *Molec. Cell Biol.* **8**, 441–451 (1988).
23. Croce, C. M. *Cell* **49**, 155–156 (1987).
24. Neuberger, M. S. & Calabi, F. *Nature* **305**, 240–243 (1983).
25. Jones, N. C., Rigby, P. W. J. & Ziff, E. B. *Genes Dev.* **2**, 267–281 (1988).

ACKNOWLEDGEMENTS. We thank the Swedish MRC, the Beit Memorial Trust and the Leukemia Society of America for support to S.P., G.P.C. and M.B. respectively.

A single origin of phenylketonuria in Yemenite Jews

Smadar Avigad*, Bernard E. Cohen†, Sofia Bauer‡, Gerard Schwartz†, Moshe Frydman*†, Savio L. C. Woo§, Yehuda Niny|| & Yosef Shiloah*¶

* Department of Human Genetics, Sackler School of Medicine, and
 || Department of Jewish History, Tel Aviv University, Tel Aviv 69978, Israel
 † Institute for Child Development, Chaim Sheba Medical Center, Tel Hashomer, Israel
 ‡ Department of Pediatrics, Hasharon Hospital, Petach Tikva, Israel
 § Department of Cell Biology, Howard Hughes Medical Institute at Baylor College of Medicine, Houston, Texas 77030, USA

PHENYLKETONURIA (PKU) is a metabolic disease caused by recessive mutations of the gene encoding the hepatic enzyme phenylalanine hydroxylase (PAH). The incidence of PKU varies widely across different geographic areas, and is highest (about 1 in 5,000 live births) in Ireland and western Scotland, and among Yemenite Jews. A limited number of point mutations account for most of the PKU cases in the European population. Here we report that a single molecular defect—a deletion spanning the third exon of the PAH gene—is responsible for all the PKU cases among the Yemenite Jews. Examination of a random sample of Yemenite Jews using a molecular probe that detects the carriers of this deletion indicated a high frequency of the defective gene in this community. Although the deleted PAH gene was traced to 25

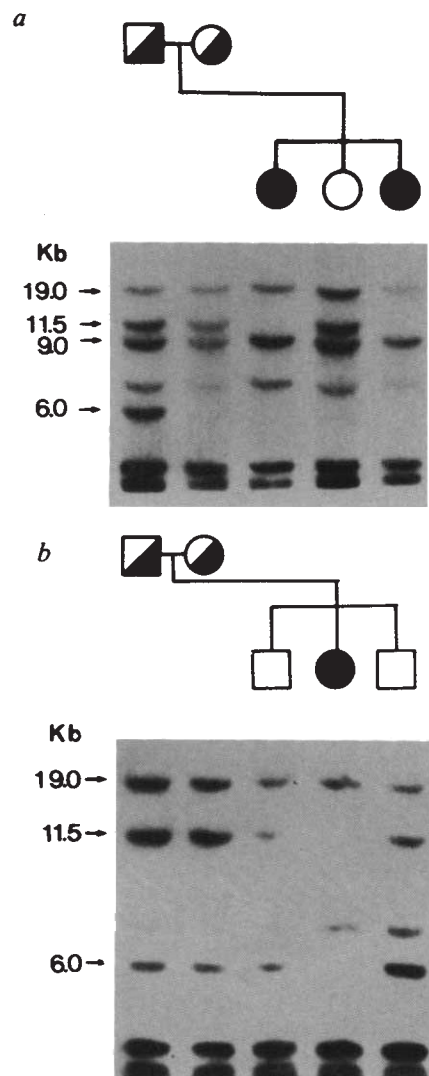


FIG. 1 Segregation of the *PvuII*(a) and *PvuII*(b) RFLPs in a non-Yemenite Jewish family (a) and a Jewish Yemenite family (b) with PKU. c, The location of the two RFLPs relative to exons 1–4 of the *PAH* gene. Normal segregation of the *PvuII*(a) alleles (6.0 and 19.0 kb) is observed in both families, whereas none of the *PvuII*(b) alleles (9.0 and 11.5 kb) are present in the DNA of the Yemenite patient. Southern blotting, hybridization and autoradiography were performed as previously described²⁰. The cDNA probe pHAN247 (ref. 5) was labelled using the random priming method²¹. Asterisks denote polymorphic restriction sites. Symbols: square, male; circle, female; filled symbol, PKU patient; half-filled symbol, PKU carrier.

different locations throughout Yemen, family histories and official documents of the Yemenite Jewish community showed that the common ancestor of all the carriers of this genetic defect lived in San'a, the capital of Yemen, before the eighteenth century.

Without PAH activity in the liver, the hydroxylation of the amino acid phenylalanine to tyrosine cannot occur, and phenylalanine accumulates in the blood and is degraded at an excessive

¶ To whom correspondence should be addressed.

rate to phenylketones. Severe mental retardation develops unless the affected child is kept on a low-phenylalanine diet shortly after birth¹. The mutations causing PKU in the Jewish population of Israel originated in 10 different communities in North Africa and Asia; Yemenite Jews are predominant among the patients, however, accounting for 25% of the affected population²⁻⁴. Because the Yemenite Jews were an ethnic isolate for many centuries, we hypothesized that PKU in this community is the result of a single mutation. We tested this hypothesis by molecularly characterizing the PAH genes of Jewish Yemenite patients in Israel, where most of this community lives today.

The full complementary DNA of the PAH gene detected 10 different restriction-fragment length polymorphisms (RFLPs), nine of which have been used for the construction of RFLP haplotypes in this gene⁵. Most of the PKU mutations identified to date in Europe and North America are associated with specific RFLP haplotypes⁶⁻¹¹. If our hypothesis is correct, then all the Jewish Yemenite patients should be homozygous for the same single haplotype. This is indeed the case in all of the 22 Yemenite PKU families identified in Israel.

The haplotype for which the patients in these families are homozygous is different from any other haplotype identified in the PAH locus so far, in that the *PvuII* (b) RFLP is missing (Fig. 1). The genomic sequence that accounts for the appearance of this RFLP, when a cDNA is used as a probe, is the third exon of the PAH gene (Fig. 1). The absence of either allele of the RFLP from the DNA of the patients could result from homozygosity for a deletion spanning this exon. We confirmed this assumption by Southern-blot analysis, in which we used

several other restriction enzymes. Constant hybridization bands, which normally characterize the PAH gene, were consistently missing from the DNA of all the Jewish Yemenite patients, indicating homozygosity of all of them for the same deletion (data not shown). We estimated the size of the deletion by hybridizing the entire cosmid cPAH10 (ref. 12) to Southern blots made with DNA from the patients. This cosmid contains 33 kilobases (kb) of genomic DNA, spanning exons 2 and 3 of the PAH gene. Figure 2 shows the absence of two *SacI* fragments, of 7.2 and 3.0 kb, and the appearance of a junction fragment of 3.5 kb, in the DNA of a Jewish Yemenite patient. This result indicates that 6.7 kb is missing from the PAH gene in this individual.

Because the deletion of exon 3 is the only molecular defect responsible for PKU in the Jewish Yemenite community, any genomic fragment spanning one of the breakpoints of the deletion (and therefore also the junction) could serve as a molecular probe for the detection of PKU carriers in this population. Heterozygotes for the deletion should be readily identifiable by the presence of a junction fragment in their DNA (Fig. 2). Carrier detection is important for genetic counselling, and is also necessary for accurate assessment of the frequency of the PKU allele in this population. For convenience, rather than using the entire cosmid cPAH10, we isolated a unique 2.6-kb *HindIII*-*EcoRI* fragment of the cosmid (designated C2.6), which is located 5' of exon 3 and contains the 5' breakpoint of the deletion. This probe clearly detected the junction of the deletion (Fig. 3). Using this probe we found that in two additional PKU families, in which only one of the parents was Yemenite, the defective PAH allele of the Yemenite parent was again the one with the exon-3 deletion. In a pilot study of the Jewish Yemenite population, we identified five carriers of the deletion among 200 randomly chosen volunteers from this community who were not related to the PKU families. Although this preliminary sample is small, it indicates a relatively high frequency of the PKU allele in this population.

To establish the origin of the deleted allele, we determined the carrier status of living grandparents in the families of the patients and random carriers by using the C2.6 probe, and questioned several members of each family about their family

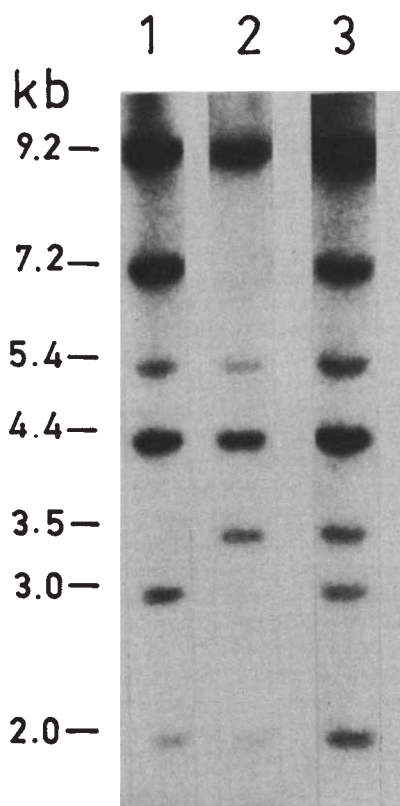


FIG. 2 Determination of the size of the deletion in the defective PAH allele. After radioactive labelling, the recombinant cosmid cPAH10 (ref. 12) was annealed to sonicated human DNA (2 mg ml^{-1}) for 30 min at 65°C to block repetitive sequences in the insert²², and then hybridized to a Southern blot prepared from *SacI*-digested DNA samples. Lane 1, control; lane 2, a Jewish Yemenite PKU patient; lane 3, the patient's father. The 7.2-kb and 3.0-kb bands are missing from the patient's DNA, and a junction fragment of 3.5 kb is detected in the DNA of the patient and his father.

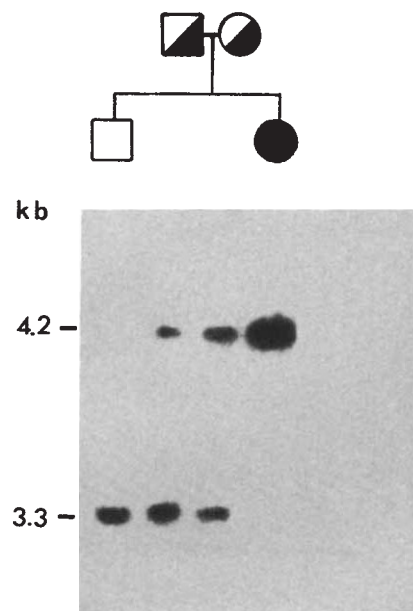


FIG. 3 Detection of a *HindIII* junction fragment of 4.2 kb in a Jewish Yemenite PKU family, using the probe C2.6 (see text). Only the 3.3-kb band is normally detected by this probe in human DNA. Southern-blot analysis was performed as indicated in the legend of Fig. 1.

history in Yemen. In this way, we traced the deletion to 25 different locations, widely spread throughout Yemen. We obtained additional information about these families from the Records of the Jewish Court in Yemen, and the Register of the Jewish Community in San'a. These two official documents of the Jewish Yemenite community list mainly marriages, divorces and trials, and provide valuable information about the life of Jews in Yemen. The information gathered from these sources showed that nine families originating in the northern part of Yemen are in fact descendants of two families from the capital, San'a. In 1809, economic hardships and religious persecution drove these two families out of San'a to three towns in the far north of the country (Fig. 4, solid arrows), from where they later spread to an additional four northern towns (Fig. 4, dotted arrows). Another 20 families came from 17 other locations in Yemen, but all originated from San'a. The migration of these families from the capital began in 1762, and continued throughout the nineteenth century and the first half of the twentieth century (Fig. 4, broken arrows). We assume, therefore, that the defective PAH allele first appeared in an individual who lived in San'a before the beginning of the eighteenth century. We believe that the deleted allele originated in the Jewish community, because this community was very orthodox, and intermarriage was unheard of. This isolation from the surrounding Moslem population is reflected in genetic profiles of the Yemenite Jews, which were constructed using various protein and immunological markers^{13,14}. Also, there are no reports of PKU in Yemenite Moslems.

The high frequency of the deleterious allele in the Jewish Yemenite population can be explained by the founder effect, genetic drift, or selective advantage of the heterozygotes. The founder effect and genetic drift result from random gene samp-

ling occurring in small isolated populations¹⁵. Heterozygote advantage has already been proposed to account for the prevalence of PKU in northwestern Europe¹⁶⁻¹⁹. Our findings should allow these different mechanisms to be examined by using molecular detection of the carriers in a defined population followed by epidemiological studies. □

Received 8 November 1989; accepted 22 January 1990.

1. Scriver, C. R., Kaufman, S. & Woo, S. L. C. in *The Metabolic Basis of Inherited Disease* 6th Edn 495-546 (McGraw-Hill, New York, 1989).
2. Cohen, B. E., Bodonyi, E. & Szeinberg, A. *Lancet* **i**, 344 (1961).
3. Cohen, B. E., Szeinberg, A., Boichis, H. & Bodonyi, E. *Pediatrics* **32**, 1069-1073 (1963).
4. Cohen, B. E. et al. *Monogr. hum. Genet.* **9**, 95-101 (1978).
5. Lidsky, A. S. et al. *Am. J. hum. Genet.* **37**, 619-634 (1985).
6. Woo, S. L. C. *Biochemistry* **28**, 1-7 (1989).
7. Scriver, C. R., Kaufman, S. & Woo, S. L. C. *An. Rev. Genet.* **22**, 301-321 (1988).
8. Abadie, V. et al. *Genomics* **5**, 936-936 (1989).
9. Daiger, S. P. et al. *Am. J. hum. Genet.* **45**, 310-318 (1989).
10. Daiger, S. P. et al. *Am. J. hum. Genet.* **45**, 319-324 (1989).
11. Okano, Y., Wang, T., Eisensmith, R. C., Gitzelmann, R. & Woo, S. L. C. *Am. J. Hum. Genet.* (in the press).
12. DiLella, A. G., Kwok, S. C. M., Ledley, F. D., Marvit, J. & Woo, S. L. C. *Biochemistry* **25**, 743-749 (1986).
13. Bodmer, J. et al. in *Histocompatibility Testing*, 125-131 (Munksgaard, Copenhagen, 1972).
14. Karlin, S., Kenett, R. & Bonne-Tamir, B. *Am. J. hum. Genet.* **31**, 341-365 (1979).
15. Cavalli-Sforza, L. L. & Bodmer, W. F. *The Genetics of Human Populations* (Freeman, San Francisco, 1971).
16. Woolf, L. I., McBean, S., Woolf, F. M. & Cahalane, S. F. *Ann. hum. Genet.* **38**, 461-469 (1975).
17. Woolf, L. I. *J. ir. med. Ass.* **69**, 398-401 (1976).
18. Woolf, L. I. *J. inher. metab. Dis.* **1**, 101 (1978).
19. Woolf, L. I. *Am. J. hum. Genet.* **38**, 773-774 (1986).
20. Shiloh, Y. et al. *Cancer Res.* **46**, 5297-5301 (1986).
21. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
22. Litt, M., & White, R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6206-6210 (1985).

ACKNOWLEDGEMENTS. We thank the staff of the Institute of Child Development at the Sheba Medical Center and the patients, volunteers and families for their patience and cooperation. We also thank Dr Tamar Schaap for reading the manuscript. S.L.C.W. is an investigator of the Howard Hughes Medical Institute. This study was supported by the Ministry of Health, Government of Israel, and the United States/Israel Binational Science Foundation.

Gene targeting in normal and amplified cell lines

Hui Zheng & John H. Wilson*

Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030, USA

TARGETED recombination in mammalian cells is rare compared with non-homologous integration¹⁻⁵. In *Saccharomyces cerevisiae* the reverse is true^{6,7}. Differences in targeting efficiency could arise because a target of unique DNA is 200 times more dilute in mammalian genomes than it is in yeast. We tested this possibility by measuring gene targeting in normal CHO cells with two copies of the dihydrofolate reductase (DHFR) gene and in amplified CHO 400 cells, which carry 800 copies⁸. If the concentration of the target gene is critical, amplified cells should show an enhanced frequency of targeted recombination relative to non-homologous integration. Using a positive/negative selection protocol³, we demonstrated that the efficiency of targeting into DHFR genes is indistinguishable in normal and amplified CHO cells. As targeting does not depend on the number of targets, the search for homology is not a rate-limiting step in the mammalian pathway of gene targeting. Thus, the difference in genome size is not the basis for the different outcomes of targeting experiments in *S. cerevisiae* and mammals.

The DHFR locus in normal and amplified CHO cells offers two distinct advantages for assessing the influence of target number on the frequency of targeted recombination: (1) The parental CHO cell line and the amplified CHO 400 line derived from it allow a direct comparison of different target numbers against a constant genetic background. Such a comparison is not possible for naturally occurring repeated sequences, such

* To whom correspondence should be addressed.

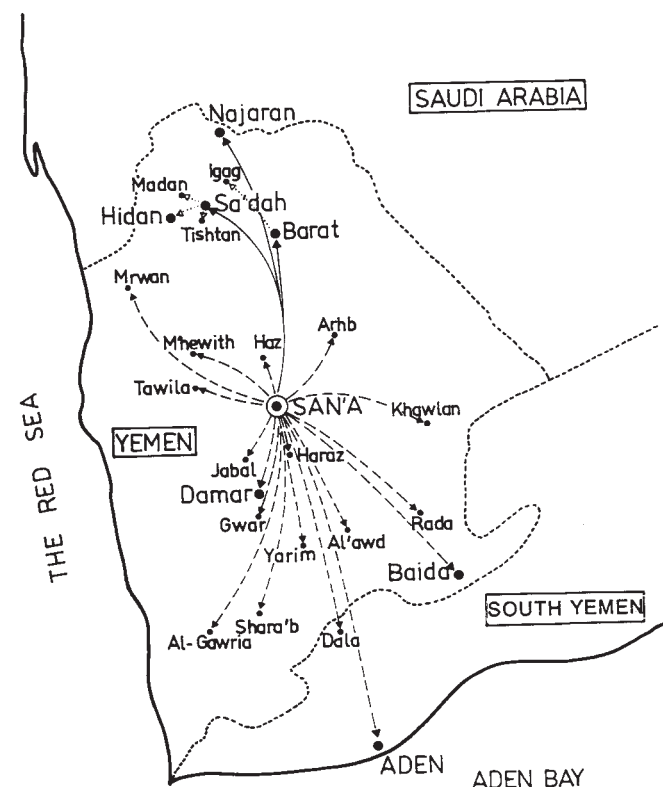


FIG. 4 Geographic spread of the PAH deletion throughout Yemen between 1762 and the early 1900s. Solid arrows, migration of two Jewish families in 1809 from San'a to three northern towns; dotted arrows, movements of their descendants within that area; broken arrows, several movements of carriers of the PAH deletion from San'a to 17 different towns and villages in other parts of the country, the first taking place in 1762.

as ribosomal RNA genes or repetitive elements, because related cell lines with different numbers of such sequences are not available; (2) the DHFR genes in the amplified cells are on standard chromosomes in a normal context. The DHFR genes are embedded in 135-kilobase (kb) amplified units with normal flanking sequences, and they seem to be expressed normally as the DHFR enzyme and messenger RNA levels are proportional to the number of amplified genes⁸.

To measure gene targeting, we used a positive/negative selection strategy³. The targeting vector contained a 4.6-kb segment of the DHFR gene, which was interrupted with a neomycin-resistance (*neo*) gene and carried the herpes simplex virus (HSV) thymidine kinase (*tk*) gene at one end (Fig. 1). Growth in G418 medium (Gibco) selects for cells in which the *neo* gene has been incorporated either by targeted recombination or non-homologous integration. Growth in FIAU (structure shown in Fig. 2) selects against cells that have incorporated the HSV-*tk* gene non-homologously through the ends of the vector^{9,10}. Targeted recombinants, which are *neo*⁺ and HSV-*tk*⁻, should be resistant to both G418 and FIAU.

The efficacy of FIAU selection was tested by measuring the effect of FIAU on the plating of CHO and CHOC 400 cells and their HSV-*tk*⁺ counterparts (Fig. 2). CHO and CHOC 400 cells were not affected by up to 5 μ M FIAU, indicating that the cellular thymidine kinase was resistant to FIAU. By contrast, cells that expressed the viral thymidine kinase were completely killed by 0.5 μ M FIAU.

To measure the frequency of gene targeting, the vector was electroporated into normal and amplified cells, which were then exposed to either G418 or to G418 plus FIAU. In three experiments carried out under the same conditions, FIAU selection significantly reduced the frequency of colony formation, presumably by eliminating most non-homologous integrants (Table

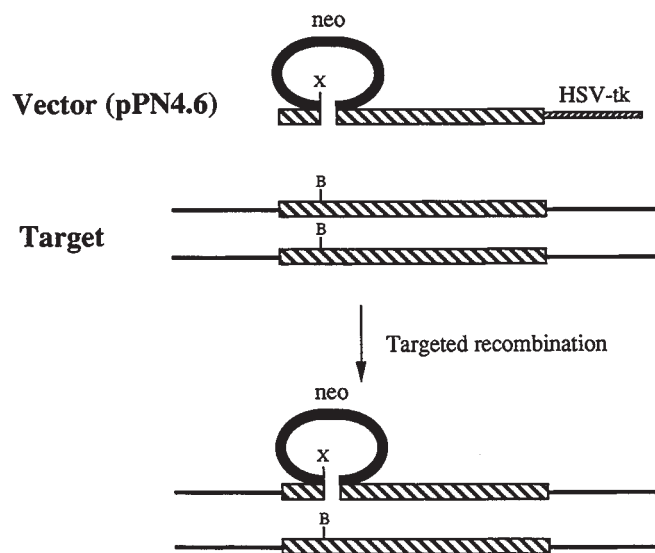


FIG. 1 Targeted recombination into the DHFR gene. The targeting vector pPN4.6 consisted of a 4.6-kb segment of the DHFR gene (large hatched rectangle), which was interrupted with a *neo* gene and carried the herpes virus *tk* gene at one end. The DHFR sequence was bounded by *Bam*HI sites and extended from intron 4 to intron 5 (ref. 22). (DHFR clones supplied by Dr L. Chasin). The neomycin-resistance (*neo*) gene from pMC1neo² was inserted at the *Bgl*II site (B) in intron 4 in a way that destroyed the *Bgl*II site and introduced a *Xba*I site (X) at one end of the *neo* cassette. Insertion of the *neo* gene separated the DHFR sequence into a 0.7-kb arm and a 3.9-kb arm. The herpes virus *tk* gene contained the same promoter and enhancer as the *neo* gene (supplied by Dr A. Bradley) and was attached so that it was transcribed in the same orientation as the *neo* gene and the DHFR gene. Targeted recombination was expected to result in the insertion of the *neo* gene into one target allele and the loss of the HSV-*tk* sequence. The other allele was expected to remain unchanged.

1). Several independent *neo*⁺*tk*⁻ clones from each experiment were analysed by Southern blotting to determine the frequency of targeted recombinants in the enriched populations. Examples of targeted recombinants from experiment 1 are illustrated in Fig. 3, with diagrams showing the diagnostic restriction fragments. In the three experiments, the frequency of targeted recombinants ranged from 30% to 100% of the *neo*⁺*tk*⁻ clones that were tested (Table 1). Comparisons of targeted recombinants relative to total integrants in CHO 400 cells versus CHO cells gave values of 2.1, 1.5, and 1.2 in the three experiments (Table 1). Thus the frequency of gene targeting in amplified genes is not significantly different from that in the parental CHO cell line.

As cells with a 400-fold difference in the target number were targeted at about the same frequency, it is unlikely that the search for homology between the vector DNA and the chromosome is a rate-limiting step in the pathway of targeted recombination in mammalian cells. If the search for homology is not rate-limiting, the difference in genome size cannot be responsible for the difference in targeting efficiency between yeast and mammalian cells. These results confirm suggestions made by others based on: (1) similar small numbers of targeted recombinants in microinjection experiments that targeted DNA into chromosomal plasmids present at 1, 4, or 5 copies per cell¹¹; and (2) the failure to detect targeted recombination into repeated

TABLE 1 Comparison of gene targeting frequencies in CHO and CHOC 400 cells

	I		II		III	
	CHO	C 400*	CHO	C 400*	CHO	C 400*
† <i>neo</i> ⁺	1,340	500	1,330	2,000	3,780	2,160
‡ <i>neo</i> ⁺ <i>tk</i> ⁻	19	36	4	31	68	54
§ Targeted	7/9	5/15	4/4	3/10	10/18	8/17
Targeted/total	1/90	1/42	1/330	1/220	1/100	1/85
¶ C 400/CHO	2.1		1.5		1.2	

Input DNA was digested with *Xho*I, which cleaves the plasmid sequences away from the vector sequences. Digested DNA was introduced into CHO and CHOC 400 cells by electroporation, using a Bio-Rad Gene Pulser apparatus. Cells (8×10^6) were resuspended in 800 μ l HBES buffer (20 mM HEPES, pH 7.05, 137 mM NaCl, 5 mM KCl, 0.7 mM Na₂HPO₄, 6 mM dextrose) containing 20 μ g DNA and electroporated at 25 μ F and 500 V per 4 mm. The treated cells were plated onto 20 100-mm Petri dishes; 48 h after plating, two plates were selected in 375 μ g ml⁻¹ G418 (Sigma) and 18 plates were selected in G418 plus 0.5 μ M FIAU. After 14 d of selection, the plates that were under G418 selection only were stained to determine the total recombinants. Independent colonies from plates exposed to both G418 and FIAU were picked (only one colony was picked from each plate), grown up, and analysed by Southern blotting (Fig. 3). The rest of the double-resistant colonies were stained and counted. Experiments were repeated three times, and are individually labelled I, II and III.

* CHOC 400 cells.

† Total number of G418-resistant colonies. From left to right, the absolute values for the frequencies of *neo*⁺ colonies relative to the number of treated cells are 1.9×10^{-4} , 0.7×10^{-4} , 1.8×10^{-4} , 2.8×10^{-4} , 5.2×10^{-4} and 3.0×10^{-4} .

‡ Total number of colonies resistant to both G418 and FIAU. These colonies should be enriched for targeted recombinants but may still include non-homologous integrants that have arisen in such a way that the *tk* gene is lost or inactive.

§ Ratio of targeted recombinants identified by Southern blotting to the total number of *neo*⁺*tk*⁻ clones analysed. From left to right, the absolute frequencies of targeted recombinants relative to the number of treated cells are 2.1×10^{-6} , 1.6×10^{-6} , 0.6×10^{-6} , 1.3×10^{-6} , 5.2×10^{-6} , and 3.5×10^{-6} .

|| The ratio of targeted recombinants to the total number of integration events. The ratio is equal to the number of *neo*⁺*tk*⁻ clones (row 2) divided by the number of *neo*⁺ clones (row 1), multiplied by the fraction of *neo*⁺*tk*⁻ clones that were targeted, as estimated from Southern blotting (row 3).

¶ The frequency of targeted recombination in CHOC 400 cells relative to that in CHO cells. For example in experiment I, this ratio is 1/42 divided by 1/90, which is 2.1.

loci, including rRNA genes at ~ 200 copies per cell¹², amplified adenosine deaminase genes at $\sim 5,000$ copies per cell (H.Z. and J.H.W., unpublished results), and repetitive elements at $\sim 100,000$ copies per cell¹³. The absence of target number dependence in mammalian cells contrasts with *S. cerevisiae*, where targeting into the repetitive rRNA genes, which are present at 140 copies per genome, was shown to be 100- to 200-fold more efficient than targeting into the unique *leu2* locus^{14,15}.

Two other observations on the effects of homology place constraints on the limiting step in the pathway of targeted recombination in mammalian cells, but do not define it. First, the frequency of gene targeting does not respond to the number of copies of the target vector microinjected into cell nuclei, suggesting that initial binding of the vector to the recombination machinery is not rate-limiting^{11,16}. Second, the frequency of gene targeting increases dramatically with increasing homology on the target vector, indicating that the rate-limiting step is very sensitive to length of homology¹⁷. These constraints are compatible with a limiting step either before or after the search for homology. For example, after loading the vector, the recombination machinery could be activated for targeting in a way that depends on the length of homology. Alternatively, the efficiency of formation or resolution of a recombination intermediate (such as a Holliday junction) could be rate-limiting and dependent on the length of homology.

In combination with the development of murine embryonic stem cells¹⁸, gene targeting in mammalian cells promises to be a powerful tool for studying development in mice and for creating animal models of human disease^{2,3,5,19}. Unfortunately, the low frequency of targeted recombination relative to non-homologous integration in mammalian cells makes gene targeting more difficult than in yeast. Identifying and overcoming the rate-limiting step would improve the usefulness of gene targeting in mammalian cells. Our experiments suggest that the search for homology is not a rate-limiting step in mammalian cells, although it may be in yeast¹⁵. Gene targeting in yeast also differs

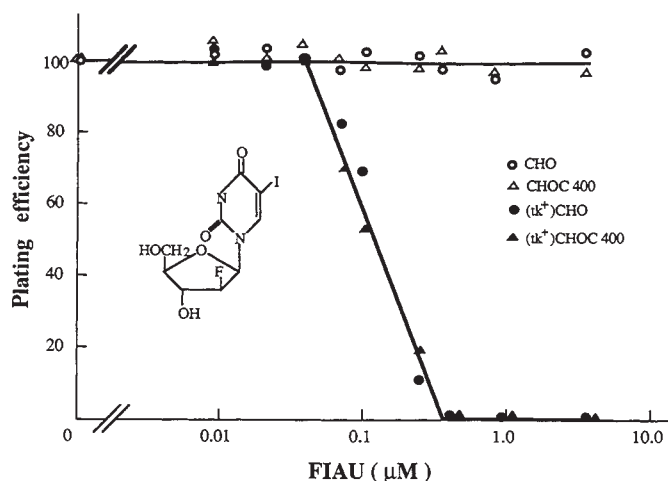
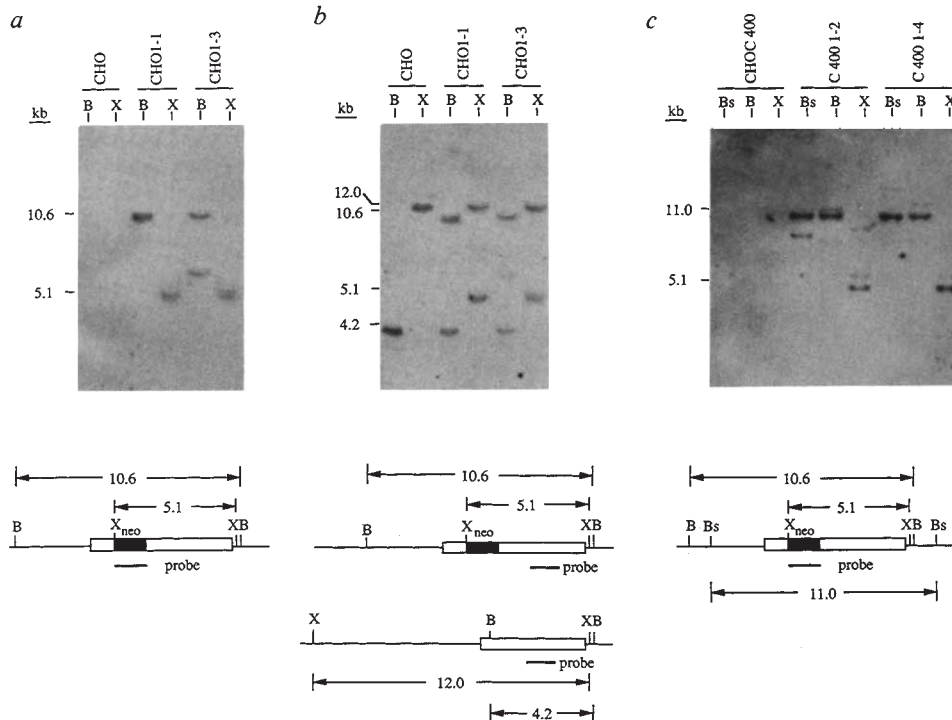


FIG. 2 The effect of FIAU on the plating efficiency of HSV-*tk*⁻ and HSV-*tk*⁺ CHO cells. FIAU, whose structure is shown, was a gift from Bristol Myers. Normal CHO cells and the amplified CHO C400 cells were provided by Dr J. Hamlin⁸. The corresponding HSV-*tk* transformed cells, (tk⁺)CHO and (tk⁺)CHO C400, were generated by introducing pPN4.6 into CHO and CHO C400 cells by electroporation and then selecting for G418-resistant cells. Individual G418-resistant colonies were subcultured to give (tk⁺)CHO and (tk⁺)CHO C400. These cells, in parallel with the wild-type CHO and CHO C400 cells were plated at 2,000 cells per 100 mm dish and incubated in Dulbecco's modified Eagle's medium, 10% fetal bovine serum, non-essential amino acids, and increasing amounts of FIAU. Colonies were stained with Coomassie brilliant blue G after 12 d and counted. Plating efficiency in the absence of FIAU addition was defined as 100%. Three different HSV-*tk*⁺ clones from each cell line were tested; all gave the same responses as the two that are shown here.

FIG. 3 Southern blot analysis of typical targeted *neo*⁺*tk*⁻ clones. *a*, DNA samples from CHO cells and two *neo*⁺*tk*⁻ CHO clones from experiment 1 (CHO1-1 and CHO1-3) were digested with *Bgl*II (B) and *Xba*I (X) and the restriction fragments were separated by electrophoresis on 0.8% agarose gels. The DNA was transferred to Gene-Screen Plus membranes and then hybridized with a *neo*-specific probe, illustrated in the drawing below the gels. The *neo*-specific probe detects only the input DNA. Because the *Bgl*II site was destroyed and a *Xba*I site was introduced by *neo* insertion, targeted recombination was expected to generate a 10.6-kb *Bgl*II fragment and 5.1-kb *Xba*I fragment. The extra bands in some of the lanes presumably resulted from non-homologous integration of a portion of the target vector, but we have not characterized them further. *b*, Same blot as in *a*, but stripped and rehybridized with the DHFR probe shown. The DHFR probe detects both the parental allele (a 4.2-kb *Bgl*II fragment and a 12.0-kb *Xba*I fragment) and the targeted allele. Comparisons of the intensities of hybridization of the 4.2-kb and 12.0-kb bands in parental CHO cells and in targeted recombinants suggests that a single allele was targeted in all recombinants. *c*, DNA samples from CHO C400 cells and two *neo*⁺*tk*⁻ CHO C400 clones were digested with *Bst*ElI (Bs) and *Xba*I and analysed by blot hybridization against the *neo* probe. *Bst*ElI digestion should generate



an 11-kb fragment which is diagnostic for the targeted allele. These samples were not hybridized against the DHFR-specific probe because the highly amplified DHFR gene obscures the diagnostic bands. Some of the faint bands resulted from partial digestion, which was apparent from other blots.

from mammalian cells in other ways. For example, targeting in yeast shows a more linear dependence on length of homology than in mammalian cells^{15,20}, and homology at the ends of the target vector is not necessary in mammalian cells³ but seems to be crucial in yeast²¹. Collectively, these differences urge caution in extrapolating from *S. cerevisiae* and reinforce the need for systematic studies in mammalian cells. □

Received 24 November 1989; accepted 11 January 1990.

- Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. *Nature* **317**, 230–234 (1985).
- Thomas, K. R. & Capecchi, M. R. *Cell* **51**, 503–512 (1987).
- Mansour, S. L., Thomas, K. R. & Capecchi, M. R. *Nature* **336**, 348–352 (1988).
- Adair, G. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 4574–4578 (1989).
- Joyner, A. L., Skarnes, W. C. & Rossant, J. *Nature* **338**, 153–156 (1989).
- Hinnen, A., Hicks, J. B. & Fink, G. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1929–1933 (1978).
- Rothstein, R. *Meth. Enzym.* **101**, 202–211 (1983).
- Milbrandt, J. D., Heintz, N. H., White, W. C., Rothman, S. M. & Hamlin, J. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6043–6047 (1981).
- Borrelli, E., Heyman, R., Hsi, M. & Evans, R. M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7572–7576 (1988).
- Roth, D. B., Porter, T. N. & Wilson, J. H. *Molec. cell. Biol.* **5**, 2299–2607 (1985).
- Thomas, K. R., Folger, K. R. & Capecchi, M. R. *Cell* **44**, 419–428 (1986).
- Steele, R. E., Bakken, A. H. & Reeder, R. H. *Molec. cell. Biol.* **4**, 576–582 (1984).
- Wallenburg, J. C., Nepveu, A. & Chartrand, P. *Nucleic Acids Res.* **15**, 7849–7863 (1987).
- Schweizer, E. C., MacKechnie, C. & Halvorson, H. O. *J. molec. Biol.* **40**, 261–277 (1969).
- Szostak, J. W. & Wu, R. *Plasmid* **2**, 536–554 (1979).
- Rommerskirch, W., Graeber, I., Grassmann, M. & Grassmann, A. *Nucleic Acids Res.* **16**, 941–952 (1988).
- Capecchi, M. R. *Science* **244**, 1288–1292 (1989).
- Bradley, A., Evans, M., Kaufman, M. H. & Robertson, E. *Nature* **309**, 255–256 (1984).
- Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. & Melton, D. W. *Cell* **56**, 313–321 (1989).
- Smolik-Utlaut, S. & Petes, T. D. *Molec. cell. Biol.* **3**, 1204–1211 (1983).
- Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6354–6358 (1981).
- Carothers, A. M., Urlaub, G., Ellis, N. & Chasin, L. A. *Nucleic Acids Res.* **11**, 1997–2012 (1983).

ACKNOWLEDGEMENTS. We thank Kathleen Marburger and De Dieu for technical assistance, and Dr Rodney Kellens, Sandra Pennington, Mark Brennenman and Xiangwei Wu for their suggestions. This work was supported by grants from the NIH and from the Robert A. Welch Foundation. H.Z. is a predoctoral fellow of the March of Dimes Birth Defects Foundation.

New data on excisions of Mu from *E. coli* MCS2 cast doubt on directed mutation hypothesis

John E. Mittler & Richard E. Lenski

Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92717, USA

ACCORDING to the directed mutation hypothesis, certain mutations in bacteria occur more frequently in environments in which the resulting phenotype is selectively favoured than in non-selective environments^{1–4}. This hypothesis therefore challenges the fundamental tenet that mutations occur spontaneously, irrespective of effects on the organism's fitness^{5–9}. One purported case of directed mutation is the excision of a Mu sequence from *Escherichia coli* strain MCS2 in minimal lactose–arabinose medium^{1,2}. Here, we show that this case can be more simply explained by an accelerated rate of excision mutation in response to non-specific physiological stresses of starvation and by slight growth of MCS2 on minimal lactose–arabinose medium.

We used the same *E. coli* MCS2 strain as in refs 1 and 2. MCS2 is a genetically manipulated strain in which the *araC* regulatory gene is placed upstream of the *lacZY* structural genes, but is separated from them by Mu bacteriophage DNA (Mu_{cts62} prophage)². If an excision of Mu produces a suitable reading frame, then the mutant cell can grow on lactose, provided that arabinose is also present as an inducer^{1,2}. (J. A. Shapiro (personal communication) has shown that a Mu-encoded function is essential for most of these excisions.) Following the notation of Cairns *et al.*¹, we refer to excision mutants as Lac(Ara)⁺, and to their MCS2 progenitor as Lac(Ara)[–].

Shapiro² observed that when Lac(Ara)[–] cells were spread on minimal agar plates containing lactose plus arabinose and incu-

bated at 32 °C, no Lac(Ara)⁺ colonies appeared on the plates for five days; after five days, Lac(Ara)⁺ colonies began to accumulate, eventually yielding plates with more than 100 colonies. A similar pattern was observed at 37 °C, although the first colonies appeared after only 2–3 days. At both temperatures, the observed frequency of mutant colonies on old plates was orders of magnitude higher than expected from the baseline mutation rate observed in growing cultures^{1,2}. Because single Lac(Ara)⁺ cells usually produce readily scored colonies on minimal lactose–arabinose (MLA) plates within 1–2 days at either temperature², this delay is not a trivial consequence of arabinose induction, catabolite repression, or slow growth^{1,2}. We repeated Shapiro's experiments with similar results; our data for 37 °C are shown in Fig. 1.

Cairns *et al.*^{1,4} interpreted these data as evidence for directed mutation in response to the presence of lactose and arabinose, but other factors could account for the accumulation of Lac(Ara)⁺ mutants. In particular, the rate of Mu excisions may increase as cells starve, irrespective of the sugars as selective agents⁶. To test this idea, we spread $\sim 2.5 \times 10^8$ Lac(Ara)[–] cells onto minimal plates with no added sugars (M plates) and incubated them at 37 °C for 0–6 days. We estimated the number of Lac(Ara)⁺ mutants on these plates by counting the colonies that appeared within two days of spraying them with lactose and arabinose (Fig. 2). No mutant colonies appeared within two days on plates sprayed with sugars immediately after cells were spread. However, the number of Lac(Ara)⁺ colonies appearing within two days of spraying increased steadily as a function of the time that cells were on the plates before addition of lactose and arabinose. This indicates that an increasing number of excision mutations occur while cells are starving, even in the absence of lactose and arabinose, and it contradicts the claim that these excisions are "... genetic events that apparently occur only under conditions of selection ..."⁴.

But can this control experiment account for all the excision mutants observed when sugars are present? Comparison of Figs 1 and 2 suggests that it does not: there are many more Lac(Ara)⁺ colonies on MLA plates than on M plates incubated for the same time (including the two days after addition of lactose and arabinose). After 6 days, for example, there were on average ~ 200 Lac(Ara)⁺ colonies on MLA plates, compared with ~ 10 on M plates. However, we have assumed in this comparison of colony numbers that the densities of Lac(Ara)[–] cells at risk for mutation are the same on M and MLA plates.

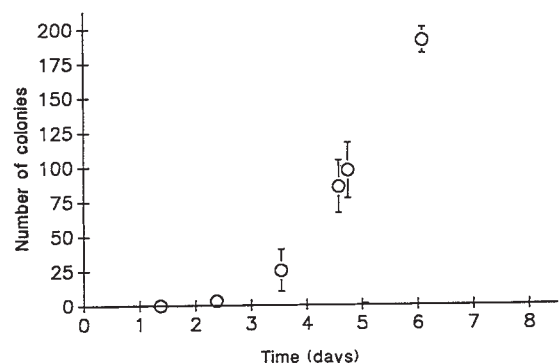


FIG. 1 The cumulative number of Lac(Ara)⁺ colonies appearing on minimal lactose–arabinose (MLA) plates incubated at 37 °C, as a function of time. About 2.5×10^8 Lac(Ara)[–] cells, grown in Davis minimal liquid¹⁰ (with 1 mg ml^{–1} glucose) for 24 h, were spread onto each of three plates. MLA plates consist of Davis minimal agar¹⁰ with 2 mg ml^{–1} each of arabinose and lactose instead of glucose. After 6 d, the number of Lac(Ara)⁺ colonies could no longer be scored unambiguously because large numbers of tiny colonies appeared in the vicinity of older Lac(Ara)⁺ colonies. Error bars represent 95% confidence intervals (*t*-distribution, *n*=3). The first point has no error bar because the standard deviation was zero; the error bar for the second point is not visible because the standard deviation was very small.

TABLE 1 Stable persistence of Lac(Ara)⁺ colony-forming cells during re-growth in fresh medium after a period of starvation

Treatment*	Log ₁₀ total viable cell density (ml ⁻¹)†	Log ₁₀ frequency of Lac(Ara) ⁺ cells‡	Proportion Lac(Ara) ⁺ visible after one day§
A. Growth in fresh medium for 1d	9.41 (0.02)	-9.70	—
B. As A, then 9d starvation	8.20 (0.02)	-5.66 (0.03)	0.12 (0.02)
C. As B, then re-growth for 1d in fresh medium¶	9.45 (0.02)	-5.63 (0.10)	0.89 (0.04)

* Populations were grown, starved, and re-grown at 32 °C in Davis minimal liquid¹⁰ (with 1 mg ml⁻¹ glucose). Means (and s.e.m.) were calculated from 10 independent replicates.

† Total viable cell density was estimated from colony counts on non-selective plates.

‡ Frequency of Lac(Ara)⁺ cells was estimated from the ratio of colony counts on MLA plates after 2d incubation at 37 °C to colony counts on non-selective plates.

§ Fraction of colonies appearing on MLA plates after 48 h (2d) incubation that were already visible after only 26.5 h (1d) incubation. Most of the Lac(Ara)⁺ colony-forming cells in the starved cultures took 2d to form visible colonies, whereas most of those that had been re-grown in fresh medium formed visible colonies after only 1d.

|| No Lac(Ara)⁺ colonies were obtained from any of 10 independent replicates. The upper limit for the frequency of Lac(Ara)⁺ cells was computed as $(ndv)^{-1}$, where n is the number of replicate populations (10), d is the mean total viable cell density (2.6×10^9 ml⁻¹), and v is the sample volume tested on MLA plates (0.2 ml).

¶ One day of re-growth corresponds to ~10 generations of binary fission, based on a 50-fold dilution factor and a 17.6-fold difference in mean total viable cell density between treatments B and C: $\log_2(50 \times 17.6) = 9.8$.

To investigate whether the presence of lactose and arabinose affected the population dynamics of Lac(Ara)⁻ cells, we sampled 'plugs' from a series of MLA and M plates incubated for different times. MLA plates had significantly more Lac(Ara)⁻ cells than the M plates, even after one day (Fig. 3), and after several days the densities differed by at least an order of magnitude. Growth by Lac(Ara)⁻ cells on plates supplemented with lactose and arabinose apparently offsets cell death, which produces a sharp decline on plates without sugars. We suggest that the higher density of Lac(Ara)⁻ cells on MLA plates was initially due to slight growth on lactose and arabinose, or on trace contaminants such as glucose. The increased difference after several days reflects cross-feeding by Lac(Ara)⁻ cells on metabolites released from growing Lac(Ara)⁺ colonies, which appear in increasing numbers on MLA plates. Indeed, we often observed 'haloes' of background growth around Lac(Ara)⁺ colonies on old MLA plates; and in minimal lactose-plus-arabinose liquid medium, Lac(Ara)⁻ cells grew significantly faster when Lac(Ara)⁺ cells were also present (data not shown).

These pronounced differences in the densities of Lac(Ara)⁻ cells are sufficient to explain the differences in the total numbers

of Lac(Ara)⁺ colonies that accumulate on MLA and M plates (Figs 1 and 2). To verify this, we crudely estimated the rate of mutation to Lac(Ara)⁺ per viable cell per day, as described in the legend to Fig. 4. Note the striking increase with time in the inferred mutation rate on both M and MLA plates; moreover, these rates are comparable. There is some indication that excision mutations occur slightly earlier on MLA plates than on M plates, but there were fewer Lac(Ara)⁺ colonies on the M plates sprayed immediately after cell spreading than on MLA plates. Variations in sugar application or local concentrations might be responsible for these subtle effects. Also, our analysis makes two assumptions that could underestimate the mutation rate on M plates relative to MLA plates. First, we assume that Lac(Ara)⁺ mutants do not die on M plates before spraying. We found an average mortality rate of ~9% per day for Lac(Ara)⁺ cells on M plates between days 1 and 5 (data not shown), versus ~78% per day for Lac(Ara)⁻ cells under comparable conditions (Fig. 3), but as the estimated mortality rate for Lac(Ara)⁺ cells was not significantly different from zero, we have conservatively assumed no death of mutants on M plates. Second, we assume that all mutant cells form visible colonies within two days of lactose and arabinose application. Any Lac(Ara)⁺ mutants present at the time of spraying and taking more than two days to form visible colonies would not be included in our estimation of the mutation rate.

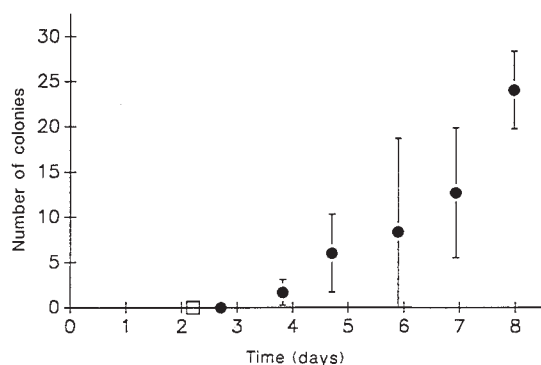


FIG. 2 The cumulative number of Lac(Ara)⁺ mutants occurring on minimal (M) plates incubated at 37 °C, as a function of time; lactose and arabinose were added to plates 2 d before Lac(Ara)⁺ colonies were scored. Day 6, for example, corresponds to a plate that sat for 4 d before addition of lactose and arabinose, and was then scored for Lac(Ara)⁺ colonies after two more days. About 2.5×10^8 Lac(Ara)⁻ cells, grown as described for Fig. 1, were spread onto each of 21 plates. Each day for 7 d (including day 0), an atomizer was used to add ~0.1 ml of a solution containing 5% lactose and 5% arabinose to 3 of these plates. The open square indicates the set of plates sprayed with lactose and arabinose immediately after cells were spread (day 0), which were counted two days later. M plates consist of Davis minimal agar without any added sugars. Error bars represent 95% confidence intervals (t -distribution, $n=3$). The first two points have no error bars because the standard deviations were zero.

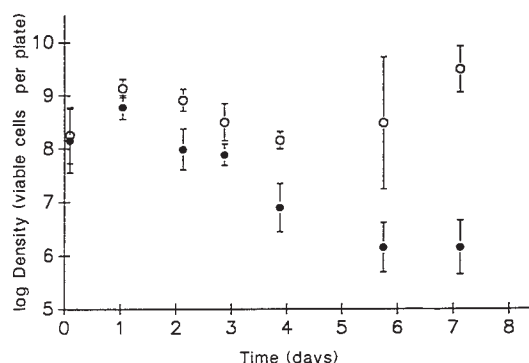


FIG. 3 Density of viable Lac(Ara)⁻ cells on MLA plates (open circles) and M plates (filled circles) incubated at 37 °C, as a function of time. About 2.5×10^8 Lac(Ara)⁻ cells, grown as described in the legend to Fig. 1, were spread onto each of three MLA and three M plates. Densities were estimated by vigorously vortexing 0.81 cm² agar plugs in saline, spreading appropriate dilutions onto lactose-arabinose tetrazolium indicator plates¹⁰, and counting the resulting Lac(Ara)⁻ colonies (which are red). After several days, many Lac(Ara)⁺ colonies (which are white) also appeared in the samples from MLA plates, but these were not counted. Error bars represent 95% confidence intervals based on the log-transformed data (t -distribution, $n=3$).

Although approximations, assumptions and sources of statistical error enter into the calculation of these mutation rates, it is clear that the rate of excision mutation per viable cell per day increases by orders of magnitude as cells sit starving for several days, irrespective of whether lactose and arabinose are present. We conclude that there is no compelling evidence for an increase in the mutation rate when the Lac(Ara)⁺ phenotype is selectively favoured, beyond that caused by the non-specific stresses of starvation.

Of course, the possibility can never be excluded that some arbitrarily small fraction of mutants is the result of a directed process. Moreover, it is often possible to produce another hypothesis which is consistent with observations that are themselves inconsistent with the earlier hypothesis: for example, the stress of starvation might have predisposed Lac(Ara)⁻ cells to directed gene mutation when exposed to lactose and arabinose. To test this *ad hoc* variant of the directed mutation hypothesis, we monitored the frequency of Lac(Ara)⁺ cells in grown, starved, and re-grown liquid cultures at 32 °C (Table 1). In freshly grown cultures, the frequency of Lac(Ara)⁺ cells is $\sim 10^{-10}$ or less^{1,2}. When these populations of MCS2 are starved for 9 days in glucose-limited minimal medium, the frequency of Lac(Ara)⁺ colony-forming cells increases to $\sim 10^{-6}$. When we transferred cells from these starved cultures into fresh minimal glucose liquid, where they grew for 10 generations, the Lac(Ara)⁺ colony-forming cells (most of which formed visible colonies after only one day) persisted at this increased frequency. Had the Lac(Ara)⁺ colony-forming cells been genotypically Lac(Ara)⁻ cells that were physiologically predisposed to directed mutation, then the frequency of Lac(Ara)⁺ colony-forming cells should have returned to the low value characteristic of freshly grown cells. We conclude that the genetic events causing the frequency of Lac(Ara)⁺ colony-forming cells to increase by several orders of magnitude occurred before exposure to lactose

and arabinose, contrary to the *ad hoc* hypothesis we have discussed.

Although the results presented here do not support the directed mutation hypothesis, they agree with the more widely held view that physiological stresses such as starvation may increase the activity of transposable genetic elements⁶. Under what circumstances this increased activity improves the fitness of the cell, the transposable element, or both, remains an interesting question. □

Received 6 July 1989; accepted 9 January 1990.

1. Cairns, J., Overbaugh, J. & Miller, S. *Nature* **335**, 142-145 (1988).
2. Shapiro, J. A. *Molec. gen. Genet.* **194**, 79-90 (1984).
3. Hall, B. G. *Genetics* **120**, 887-897 (1988).
4. Cairns, J. *Nature* **336**, 527-528 (1988).
5. Stahl, F. W. *Nature* **335**, 112-113 (1988).
6. Symonds, N. *Nature* **337**, 119-120 (1989).
7. Partridge, L. & Morgan, M. J. *Nature* **336**, 22 (1988).
8. Charlesworth, D., Charlesworth, B. & Bull, J. J. *Nature* **336**, 525 (1988).
9. Lenski, R. E. *Trends Ecol. Evol.* **4**, 148-150 (1989).
10. Carlton, B. C. & Brown, B. J. in *Manual for Methods for General Microbiology* (ed. Gerhardt, P.) 222-242 (Am. Soc. Microbiol., Washington DC, 1981).

ACKNOWLEDGEMENTS. We thank Francisco Ayala, John Cairns, Richard Hudson, Bruce Levin, Toai Nguyen, Michael Rose and Michael Travisano for their comments, and James Shapiro for making his strains available to us and for critical comments. R.E.L. was supported by the NSF and the NIH.

Dependence of the torsional rigidity of DNA on base composition

Bryant S. Fujimoto & J. Michael Schurr

Department of Chemistry, University of Washington, Seattle, Washington 98195, USA

THE *Escherichia coli* phage 434 repressor binds as a dimer to the operator of the DNA helix. Although the centre of the operator is not in contact with protein, the repressor binding affinity can be reduced at least 50-fold by changing the sequence there¹: operators with A·T base pairs near their centre bind the repressor more strongly than do operators with G·C base pairs at the same positions. To explain these observations, it has been proposed that the base composition at the centre of the operator affects the affinity of the operator for repressor by altering the ease with which operator DNA can undergo the torsional deformation necessary for complex formation^{1,2}. In this model, the variation in binding affinity would require the torsion constant to have specific values and to change in a sequence-dependent manner¹. We have now measured torsion constants for DNAs with widely different base compositions. Our results indicate that the torsion constants depend only slightly on the overall composition, and firmly delimit the range of values for each. Even the upper-limit values are much too small to account for the observed changes in affinity of the 434 repressor. These results rule out simple models that rely on substantial generic differences in torsion constant between A·T-rich sequences and G·C-rich sequences, although they do not rule out the possibility of particular sequences having abnormal torsion constants.

Torsion constants are obtained from the time-resolved fluorescence polarization anisotropy (FPA) of intercalated ethidium dye. By monitoring the rotational motion of the transition dipole of the intercalated dye, it is possible to infer the torsion constant of the host DNA. Details of the theory and experiment are described elsewhere³⁻¹⁰. The estimated torsion constant depends on the value assumed for the dynamic bending rigidity, or dynamic persistence length P_d . A lower bound torsion constant α is obtained by assuming $P_d = \infty$, which assigns all depolarization to twisting motions. An upper bound torsion constant, $\alpha' = (1.9)\alpha$, is obtained by assuming $P_d = 500$ Å (ref. 11). The

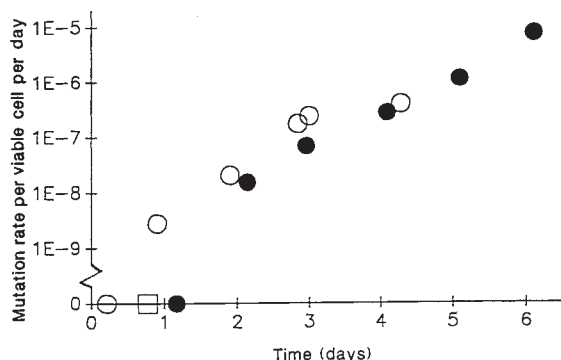


FIG. 4 Inferred rates of mutation per viable cell per day from Lac(Ara)⁻ to Lac(Ara)⁺ on MLA plates (open circles) and M plates (filled circles and open square) incubated at 37 °C, as a function of time. These rates were calculated as $u(t) = c(t')/d(t)$, where $u(t)$ is the time-dependent mutation rate, $d(t)$ is the number of Lac(Ara)⁻ cells at risk at time t , and $c(t')$ is the rate of appearance of new colonies at some later time t' . The time t' is given by $t' = t + h(t)$, where $h(t)$ is the elapsed time between occurrence of a mutation and appearance of the resulting colony. Freshly grown Lac(Ara)⁺ cells typically produce visible colonies within 1 d after plating on MLA, but starved Lac(Ara)⁺ cells usually require 2 d to form colonies (Table 1, footnote). We therefore set $h(t) = 1 + t/(1 + t)$, such that $h(0) = 1$, $h(1) = 1.5$, $h(2) = 1.67$, and so on. The precise form of this function has little effect on the overall appearance of this figure, provided that $1 < h(t) < 2$. The rate of appearance of Lac(Ara)⁺ colonies, $c(t')$, was estimated from the difference between adjacent means in Figs 1 and 2, divided by the time elapsed between them; the later time point for each pair was designated t' . The number of Lac(Ara)⁻ cells at risk for mutation, $d(t)$, was obtained by interpolation between adjacent logarithmic means in Fig. 3, to provide estimated cell densities $h(t)$ days before each estimate of $c(t')$. Solving for t requires iteration of $t' = t + h(t)$. Inferred mutation rates are not available for MLA plates after 4 d because it becomes too difficult to score the number of Lac(Ara)⁺ colonies after 6 d (Fig. 1).

actual torsion constant α_a can be estimated from the lower bound α and any assumed P_d according to Fig. 1. We emphasize that the lower bound α is proportional to the actual torsion constant α_a , and faithfully monitors any changes in the actual torsion constant. The uncertainty in the magnitude of α_a arises primarily from uncertainty regarding P_d . Torsion constants obtained using our best estimate $P_d = 1,500 \text{ \AA}$ (refs 10, 12), are in good agreement with the results of recent static ligation studies¹³, as described in the legend to Fig. 1.

Table 1 presents torsion constants for several linear DNAs and linearized plasmids. The torsion constants are uniform in the sense that the same best-fit values of α are obtained for data extending over several different time spans (see Fig. 2). This implies that any major weaknesses in the torsional rigidity must occur less often than 1 in 1,000 base pairs (bp) or more frequently than 1 in 20 bp⁵. These values are insensitive to the precise location of the intercalated ethidium molecules, because the FPA measurements probe collective motions of many (50–1,000) base pairs extending far beyond the binding site. Anyway, ethidium binds relatively nonspecifically to long DNAs^{14–16} and heptamers¹⁷, despite its preference for pyrimidine-purine sequences in tetramers and dinucleotides. Intercalated ethidium has no effect on the torsion constant even up to much higher binding ratios (1 in 5 bp)¹⁸ than those of our studies (~1 in 300 bp).

Fluorescence decay measurements on $(dA-dT)_n \cdot (dA-dT)_n$ reveal the presence of additional species besides intercalated and free dye⁸. This implies two or more modes of binding of the dye. A proper analysis of the data is precluded, unless the orientations of the transition dipole and the local librational dynamics of the dye in each binding site are known. Possible weakening of the long-range torsional rigidity by interruptions of the helix, for example at branched structures or cruciforms, is also an important concern with $(dA-dT)_n \cdot (dA-dT)_n$. In our opinion, no credible values for the torsional rigidity of continuous duplex $(dA-dT)_n \cdot (dA-dT)_n$ have been reported so far, and none appears in Table 1. The results of Millar *et al.*¹⁹ and

Ashikawa *et al.*²⁰ for unfractionated, uncharacterized, commercial synthetic DNAs are omitted from Table 1 for reasons described in the legend.

We omitted torsion constants determined by transient photodichroism (TPD)^{21,22} from Table 1. They were also neglected by Hogan and Austin¹ in favour of values deduced from the bending rigidity and an assumed Poisson ratio. The decay of the transient photodichroism of intercalated methylene blue is described by equations similar to those for FPA and is similarly interpreted³. TPD is the best method available for determining P_d . But, the precision and reliability of the torsion constants obtained from TPD measurements²² are significantly lower than those obtained from FPA measurements for several reasons. The TPD instrument response function is much wider ($\geq 15 \text{ ns}$) than that of our FPA apparatus (0.55 ns, recently improved to $\leq 60 \text{ ps}$), and is not deconvoluted from the data. For the 400-bp DNA of ref. 22, ~46% of the total depolarization due to twisting occurs in the first 20 ns, which is largely inaccessible in those measurements. TPD data from 20 to 300 ns, during which twisting predominates, are sometimes not fitted very well, because they are statistically overwhelmed by extensive data from much longer time spans, during which bending predominates. The problem arises because the theory describing the bending dynamics is still inadequate. The resulting systematic discrepancy between the assumed theory and the actual bending dynamics over the long time span can degrade the quality of the fit at early times. None of the TPD torsion constants was shown to be independent of the time span of the data fitted, as were the FPA values in Table 1. The initial anisotropies of the TPD measurements are considerably lower than those observed for FPA measurements. If this difference stems from anisotropic dye wobble, the theory assumed in the TPD data analysis would be systematically in error. TPD experiments provide no estimate of the intercalated fraction of methylene blue, and are equally sensitive to intercalated, outside-bound and free dye. The steady-state fluorescence assay used to measure the binding constant for methylene blue does not distinguish intercalation from other modes of binding. By contrast, time-resolved FPA experiments yield estimates of both intercalated and nonintercalated fractions of ethidium⁹, and are much more sensitive to the former.

Overall base composition has only a slight effect on the torsion constant over the range 34–100% G-C. This insensitivity implies that the torsion constant, on average, is insensitive to which base pairs are involved. But the larger values observed for plasmid pUC8 and pBR322 DNAs indicates that particular sequences can exert different effects on the apparent torsion constant. Sequence-dependent tilting of the ethidium-binding site, so as to reduce the angle between the transition dipole and the helix axis, would increase the apparent α even without a change in torsion constant.

The two DNAs with lowest G-C content, $\phi 29$ (34% G-C) and calf thymus (39% G-C), undoubtedly have many A-T-rich regions. If α were substantially lower in these regions, we would expect (1) the A-T-rich regions to serve as weak spots, so that α is nonuniform and decreases significantly with increasing time span of the data fitted⁵, and (2) the average torsion constant to be significantly lower than that observed for DNAs with higher G-C content. Because neither of these expectations is met, the extent to which long A-T-rich regions can differ in torsion constant from long G-C-rich regions is limited.

Studies indicate that the 434 repressor binds as a dimer to a 14-bp operator, making contact with base pairs 1–5 and 10–14, but not base pairs 6–9^{23,24}. Koudelka *et al.*² measured the binding affinity of the 434 repressor for several operators, including ACAATATATATTGT (I) and ACAATAGCTATTGT (II), where the four underlined base pairs are not in contact with protein. Operator I has a binding affinity 50 times that of operator II. Crystallographic analysis² of the complex shows that the operator DNA is slight overtwisted with an angle of 39° between base pairs 7–8, and 36° between base pairs 6–7 and 8–9.

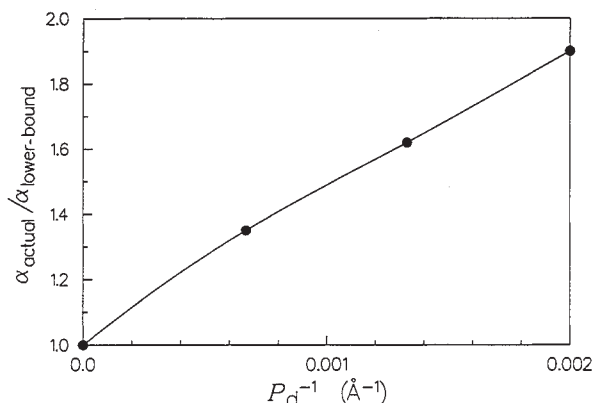


FIG. 1 Ratio of actual to lower bound torsion constant (α_a/α) versus inverse P_d , for long DNAs. Analysis of transient photodichroism data^{11,21} using Barkley-Zimm theory³⁰ and brownian dynamics simulations²⁹, and analysis¹² of transient electric dichroism data^{32,33} using a recently developed normal-mode theory³¹ indicate that the dynamic persistence length can be much (2.5 to 3-fold) larger than the static persistence length¹², which is ~500 $\text{\AA}$ for native DNAs in the present range of ionic strengths. Here, simulated data were generated using the intermediate zone twisting correlation function of Allison and Schurr^{34,30} and the Barkley-Zimm bending correlation function³⁰ with different values of P_d . These simulated data were fitted using the intermediate zone formula^{4,30} and $P_d = \infty$ to obtain a lower bound α . The ratio α_a/α as plotted provides an empirical method to estimate the actual torsion constant from the lower bound value, once P_d is known. An upper bound torsion constant $\alpha' = (1.9)\alpha$ is obtained by assuming $P_d = 500 \text{ \AA}^{-1}$. Our best estimate, $P_d = 1,500 \text{ \AA}^{-1}$ yields $\alpha_a = (1.35)\alpha$, giving $\alpha_a = 5.1, 5.8$ and $6.5 \times 10^{-12} \text{ dyn cm}$ for $\phi 29$, pUC8 and pBR322, respectively. These values agree well with recent static ligation measurements, which yield $\alpha_a = (5.9 \pm 0.6) \times 10^{-12} \text{ dyn cm}$ (ref. 13).

Koudelka *et al.* attribute the 50-fold difference in binding affinity between I and II to a greater resistance to twisting deformation of II, specifically to a larger torsion constant associated with the presence of G·C rather than A·T base pairs at positions 7 and 8². This explanation seems most unlikely for two reasons. First, we find no evidence for a strong or even moderate dependence of the torsion constant on base composition. Two of the three papers cited in support of such a proposition report only the static persistence length (or bending rigidity)^{25,26}, and the third is based on TPD measurements, which monitors primarily the dynamic bending rigidity, as noted above. It has been suggested that the bending and twisting rigidities are related through a constant Poisson ratio¹, but there is no firm empirical or theoretical basis for this. Second, the angular deformations are much too small to generate such a large binding-constant ratio, or free-energy difference, given the actual magnitude of the torsion constant. Assuming that the central twist angles adopt the canonical value of 34.6° when not bound to the repressor, the energy required to enlarge the central three twist angles from (34.6°, 34.6°, 34.6°) to (36°, 39°, 36°) is

$$\Delta E = (\alpha_0/2)[2((1.4)\pi/180)^2 + ((4.4)\pi/180)^2] \\ = (3.55 \times 10^{-3})\alpha_0 \quad (1)$$

If it is assumed for operator I that $\alpha_0 = 0$, and for operator II that α_0 equals the upper-bound $\alpha' = 7.6 \times 10^{-12}$ dyn cm for (dG-dC)_n·(dG-dC)_n, then $(\Delta E(\text{II}) - \Delta E(\text{I}))/k_B T = 0.667$ at 20 °C, where k_B is Boltzmann's constant and T is the absolute temperature. The entropy change of the torsional degrees of freedom upon binding (ΔS) is negative and smaller for a larger torsion constant. Therefore $\Delta S(\text{II}) - \Delta S(\text{I}) > 0$, and the ratio of the binding constants (K) for these two different operators with the 434 repressor is

$$K(\text{I})/K(\text{II}) = \exp [\Delta E(\text{II}) - \Delta E(\text{I}) - T(\Delta S(\text{II}) - \Delta S(\text{I}))/k_B T] \\ \leq \exp (0.667) = 1.95$$

The predicted binding-constant ratio is too small by a factor of >25. Therefore, even the maximum allowed variation in torsion constant is grossly insufficient to explain the binding data. It is conceivable that AGCT is a special sequence with an unusually large torsion constant, but to explain the binding constant ratio, the torsion constant of this sequence would have to be at least 6–7 times stiffer than the upper bound values for *Micrococcus luteus* DNA (77% G·C) and (dG-dC)_n·(dG-dC)_n, which is most unlikely.

In view of these difficulties with the proposed binding model¹, alternate explanations merit consideration. The binding discrimination could arise from sequence-dependent variations in the equilibrium structure (for example, twist) of the operator

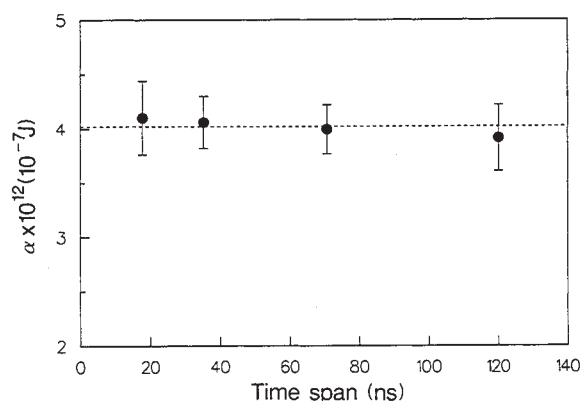


FIG. 2 Apparent torsion constant α versus time span for *M. luteus* DNA. The four points are values obtained by fitting data over the indicated time spans: 0–18, 0–36, 0–72 and 0–120 ns. The uniformity of α with time span indicates that any major torsional rigidity weaknesses must occur less often than 1 in 1,000 bp or more frequently than 1 in 20 bp.

TABLE 1 Lower bound and upper bound torsion constants for various linear DNAs

DNA	N+1 (bp)	% G·C	μ^* (m)	$\alpha \times 10^{12} \dagger$ (dyn cm)	$\alpha' \times 10^{12} \ddagger$ (dyn cm)	Ref.
$\phi 29$	17,400	34	0.01	3.8	7.2	5, u.d.
$\phi 29$	17,400	34	0.1	3.5	6.7	6, u.d.
Calf thymus	$\sim 10^4$	39	0.01	3.8	7.2	u.d.
Calf thymus	$\sim 10^4$	39	0.1	3.5	6.7	u.d.
λ	48,000	49	0.01	3.8	7.2	u.d.
<i>M. luteus</i>	$\sim 10^4$	77	0.02	4.0	7.6	u.d.
(dG-dC) _n ·(dG-dC) _n	230	100	0.02	4.0	7.6	8
(dG-dC) _n ·(dG-dC) _n	590	100	0.02	4.0	7.6	8
M13mp7§	7,238	41	0.01	3.8	7.2	7
pUC8§	2,717	51	0.1	4.3	8.2	18, u.d.
pBR322§	4,363	54	0.1	4.8	9.1	9, 18, u.d.

Results are obtained at 20 °C. Reduced χ^2 values for the anisotropy fits are almost invariably <1.2 and in many cases are <1.05. All of these torsion constants are uniform in the sense that the same best-fit values of α are obtained for data extending over several different timespans, as illustrated in Fig. 2 for *M. luteus* DNA. The FPA data are deconvoluted using the instrument response function (0.55 ns, recently improved to ≤ 60 ps). The length distributions of all samples are characterized by gel electrophoresis and for (dG-dC)_n·(dG-dC)_n, also by sedimentation. Formulae appropriate for the particular lengths were used to analyse the FPA in all cases^{3,8}. These values are robust in the following sense. At an ionic strength of 0.1 M the torsion constant of linear DNA is independent of T from 0 to 70 °C^{6,10} and intercalated chloroquine⁹ or ethidium¹⁸ up to 1 per 5 bp. It is only weakly dependent on ionic strength over the range 0.01–1.0 M^{10,19} or on bound spermidine up to neutralization of ~85% of the phosphates¹⁰. For some DNAs, such as $\phi 29$, calf thymus, plasmids pUC8 and pBR322, the same values have been repeatedly obtained for many different samples by various investigators (J. C. Thomas, J. H. Shibata, P. Wu and B.S.F.) over several years. The sample preparation procedures, experimental set-ups and data analysis programs also changed, as various improvements were progressively installed, without affecting the mean values. The results of Millar *et al.*¹⁹ and Ashikawa *et al.*²⁰ for commercial synthetic DNAs are omitted from Table 1. Both groups employed an incorrect anisotropy formula. Also, their unfractionated and uncharacterized samples undoubtedly consisted entirely, or to a large degree, of DNAs so short that the intermediate zone formula^{3–10} used in the data analysis was invalid¹¹. u.d., Unpublished data.

* μ , ionic strength.

† Lower bound torsion constants α . Estimated uncertainties are about $\pm 8\%$ for $\phi 29$, calf thymus, λ , *M. luteus* and M13mp7, $\pm 10\%$ for (dG-dC)_n·(dG-dC)_n, and $\pm 6\%$ for plasmids pUC8 and pBR322.

‡ Upper bound torsion constants $\alpha' = (1.9)\alpha$. Percentage errors as for lower bound torsion constants α .

§ Linearized.

rather than from its deformability (for example, torsion constant). If the equilibrium value of one or more of the central twist angles in unbound operator II were much smaller, then it would have to undergo a much larger deformation to bind the repressor.

The possibility that sequence dependence of the bending rigidity accounts for the observed differences in binding affinity was proposed by Hogan and Austin² and also considered by Koudelka *et al.*¹. This possibility was discounted by Koudelka *et al.* on the basis of gel retardation shift assays of wild-type and mutant repressors complexed with 82-bp restriction fragments containing the wild-type operator (I) at either the end or the centre. They attribute the gel retardation to a bend induced in the DNA by the repressor. But the slight bend of the DNA in the crystal structure of the 434 repressor/operator complex seems much too small to influence its gel mobility significantly. A comparable gel retardation shift is observed on the binding of catabolite activator protein (CAP) to a 203-bp DNA fragment²⁷, but in that case the bend is thought to be 90–180° (ref. 28). It seems that either the complex of 434 repressor with restriction fragments in solution shows a much greater bend than the reported crystal structure, or that the gel retardation shift arises primarily from some cause other than a permanent bend. We conclude that sequence dependence of either the equilibrium structure or the bending rigidity of the operator remains a possible determinant of the binding affinity of phage 434 repressor. □

Received 24 July; accepted 27 December 1989.

1. Koudelka, G. B., Harbury, P., Harrison, S. C. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4633–4637 (1988).
2. Hogan, M. E. & Austin, R. H. *Nature* **329**, 263–266.

3. Schurr, J. M. *Chem. Phys.* **84**, 71-96 (1984).
4. Allison, S. A. & Schurr, J. M. *Chem. Phys.* **41**, 35-59 (1979).
5. Thomas, J. C., Allison, S. A., Appellof, C. J. & Schurr, J. M. *Biophys. Chem.* **12**, 177-188 (1980).
6. Thomas, J. C. & Schurr, J. M. *Biochemistry* **22**, 6194-6198 (1983).
7. Shibata, J. H., Wilcoxon, J., Schurr, J. M. & Knauf, V. *Biochemistry* **23**, 1188-1194 (1984).
8. Fujimoto, B. S., Shibata, J. H., Schurr, R. L. & Schurr, J. M. *Biopolymers* **24**, 1009-1022 (1985).
9. Wu, P. *et al.* *Biochemistry* **27**, 8128-8144 (1988).
10. Schurr, J. M., Fujimoto, B. S., Wu, P. & Song, L. in *Applications of Fluorescence Spectroscopy* (ed. Lakowicz, J. R.) (in the press).
11. Shibata, J. H., Fujimoto, B. S. & Schurr, J. M. *Biopolymers* **24**, 1909-1930 (1985).
12. Song, L. & Schurr, J. M. *Biopolymers* (in the press).
13. Taylor, W. H. & Hagerman, P. J. *J. molec. Biol.* (in the press).
14. Waring, M. J. *J. molec. Biol.* **13**, 783-801 (1965).
15. Le Pecq, J. B. & Paoletti, C. *J. molec. Biol.* **27**, 87-106 (1967).
16. Müller, W. & Crothers, D. M. *Eur. J. Biochem.* **54**, 267-277 (1975).
17. Nelson, J. W. & Tinoco, I. Jr *Biopolymers* **23**, 213-233 (1984).
18. Wu, P., Fujimoto, B. S., Song, L. & Schurr, J. M. *Biophys. Chem.* (manuscript submitted).
19. Millar, D. P., Robbins, R. J. & Zewail, A. H. *J. chem. Phys.* **76**, 2080-2094 (1982).
20. Ashikawa, I., Kinosita, K. Jr. & Ikegami, A. *Biochim biophys. Acta* **789**, 87-93 (1984).
21. Hogan, M., Wang, J., Austin, R. H., Monitto, C. L. & Herschkowitz, S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3518-3522 (1982).
22. Hogan, M., LeGrange, J. & Austin, R. H. *Nature* **304**, 752-754 (1983).
23. Anderson, J. E., Ptashne, M. & Harrison, S. C. *Nature* **326**, 888-891 (1987).
24. Koudelka, G. B., Harrison, S. C. & Ptashne, M. *Nature* **326**, 886-888 (1987).
25. Thomas, T. J. & Bloomfield, V. A. *Nucleic Acids Res.* **11**, 1919-1930 (1983).
26. Chen, H. H., Rau, D. C. & Charney, D. C. *J. biomolec. Struct. Dyn.* **2**, 709-719 (1985).
27. Wu, H.-M. & Crothers, D. M. *Nature* **308**, 509-513 (1984).
28. Liu-Johnson, H.-N., Gartenberg, M. R. & Crothers, D. M. *Cell* **47**, 995-1005 (1986).
29. Allison, S. A., Austin, R. H. & Hogan, M. E. *J. chem. Phys.* **90**, 3845-3854 (1989).
30. Barkley, M. D. & Zimm, B. H. *J. chem. Phys.* **70**, 2991-3007 (1979).
31. Song, L. & Schurr, J. M. *Biopolymers* (in the press).
32. Diekmann, S., Hillen, W., Morgenmeyer, B., Wells, R. D. & Pörschke, D. *Biophys. Chem.* **15**, 263-270 (1982).
33. Pörschke, D., Zacharias, W. & Wells, R. D. *Biopolymers* **26**, 1971-1974 (1987).

ACKNOWLEDGEMENTS. This work was supported in part by the NSF.

ERRATUM

Palaeontological and isotope evidence for warm saline deep waters in Ordovician oceans

L. Bruce Railsback, Spafford C. Ackerly,
Thomas F. Anderson & John L. Cisne

Nature **343**, 156-159 (1990)

In this letter, Figs 2 and 3 have been erroneously transposed; the legends are correct. Also, the equation for oxygen isotopic composition is incorrect by a factor of 10^3 and should read

$$\delta^{18}\text{O}(\text{‰}) = [({}^{18}\text{O}/{}^{16}\text{O})_{\text{sample}}/({}^{18}\text{O}/{}^{16}\text{O})_{\text{standard}} - 1] \cdot 1,000$$

CORRECTION

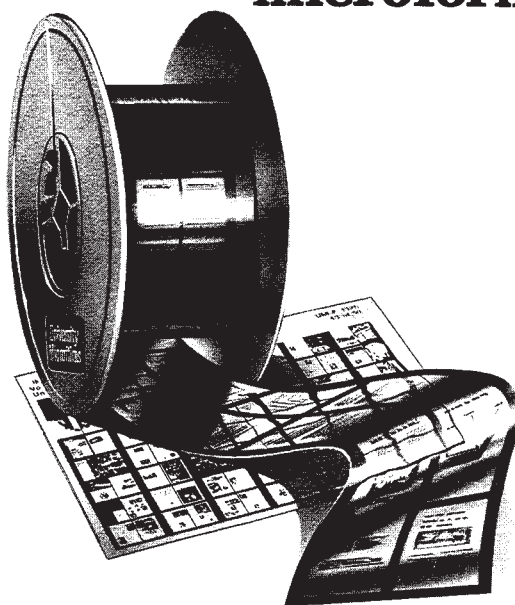
An inducible transcription factor activates expression of human immunodeficiency virus in T cells

Gary Nabel & David Baltimore

Nature **326**, 711-713 (1987)

THE authors of the above paper have recently become aware of a typographical error in the sequence reported in Fig. 1. The mutant sequence should read TCT rather than CTC at both sites and 5 base pairs upstream of the first κ B site, the T should read C. These changes do not alter the conclusions of the manuscript in any way and were found only recently upon re-sequencing the original plasmid. The CTC mutation has independently been introduced and also works identically to the TCT mutation.

nature
is available in
microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone () _____

University
Microfilms
International